LIE GROUPS; LIE ALGEBRAS

Melvin Hausner and
Jacob T. Schwartz

notes on
mathematics
and its
applications

GORDON AND BREACH

Digitized by the Internet Archive
in 2022 with funding from
Kahle/Austin Foundation

https://archive.org/details/liegroupsliealge0000melv

LIE GROUPS · LIE ALGEBRAS

Notes on Mathematics and Its Applications

General Editors: Jacob T. Schwartz, Courant Institute of Mathematical Sciences and Maurice Lévy, Université de Paris

E. Artin, ALGEBRAIC NUMBERS AND ALGEBRAIC FUNCTIONS
R. P. Boas, COLLECTED WORKS OF HIDEHIKO YANABE
M. Davis, A FIRST COURSE IN FUNCTIONAL ANALYSIS
M. Davis, LECTURES ON MODERN MATHEMATICS
J. Eells, Jr., SINGULARITIES OF SMOOTH MAPS
K. O. Friedrichs, ADVANCED ORDINARY DIFFERENTIAL EQUATIONS
K. O. Friedrichs, SPECIAL TOPICS IN FLUID DYNAMICS
M. Hausner and J. T. Schwartz, LIE GROUPS; LIE ALGEBRAS
P. Hilton, HOMOTOPY THEORY AND DUALITY
F. John, LECTURES ON ADVANCED NUMERICAL ANALYSIS
A. M. Krall, STABILITY TECHNIQUES FOR CONTINUOUS LINEAR SYSTEMS
H. Mullish, AN INTRODUCTION TO COMPUTER PROGRAMMING
J. T. Schwartz, W* ALGEBRAS
A. Silverman, EXERCISES IN FORTRAN
J. J. Stoker, NONLINEAR ELASTICITY

Additional volumes in preparation

Lie Groups
Lie Algebras

by MELVIN HAUSNER
Department of Mathematics
New York University

and

JACOB T. SCHWARTZ
Courant Institute of
Mathematical Sciences
New York University

G | B

GORDON AND BREACH
SCIENCE PUBLISHERS · NEW YORK · LONDON · PARIS

COPYRIGHT © 1968 BY GORDON AND BREACH
Science Publishers, Inc.
150 Fifth Avenue, New York, N. Y. 10011

Library of Congress Catalog Card Number 66-28064

Editorial Office for Great Britain:

Gordon and Breach Science Publishers Ltd.
8 Bloomsbury Way, London W. C. 1

Editorial Office for France:

7-9 rue Emile Dubois,
Paris 14^e

Distributed in France by:

Dunod Editeur
92 rue Bonaparte
Paris 6^e

Distributed in Canada by:

The Ryerson Press
299 Queen Street West
Toronto 2B, Ontario

512.897
H376l

Printed in Germany by VEB Offizin Andersen Nexö in Leipzig

Editor's Preface

A large number of mathematical books begin as lecture notes; but, since mathematicians are busy, and since the labor required to bring lecture notes up to the level of perfection which authors and the public demand of formally published books is very considerable, it follows that an even larger number of lecture notes make the transition to book form only after great delay or not at all. The present lecture note series aims to fill the resulting gap. It will consist of reprinted lecture notes, edited at least to a satisfactory level of completeness and intelligibility, though not necessarily to the perfection which is expected of a book. In addition to lecture notes, the series will include volumes of collected reprints of journal articles as current developments indicate, and mixed volumes including both notes and reprints.

JACOB T. SCHWARTZ
MAURICE LÉVY

METHODIST COLLEGE LIBRARY
Fayetteville, N. C.

47144

Preface

The theory of semi-simple Lie groups, the triumphant accomplishment of Lie and Killing, of E. Cartan and Weyl, is surely among the most beautiful of mathematical theories. It is a theory of the most perfect abstract type: setting forth from a broad general definition, the analysis succeeds in enumerating all its instances. The objects of the Lie theory therefore constitute the sole possible groups of symmetry of many other algebraic, analytic, geometric, and physical theories; so that the Lie theory is not merely elegant, but consequential also.

Our presentation follows standard lines. We begin with Lie groups. It soon becomes apparent that the analysis of these objects may largely be reduced to an investigation in linear algebra, i.e., to the analysis of the Lie algebra of a Lie group. Normal considerations of ideal theory then lead to the special sub-class of semi-simple algebras; in the complex case these are analyzed completely by the method of Killing and Cartan as successfully simplified by Dynkin. We present this part of the analysis making free use of simplifications afforded by analytic and topological considerations. At the end of this theory we find the four main series of classical Lie algebras, and the five exceptional algebras. The main series of algebras are of recognised importance in mathematics. As to the exceptional algebras, the convinced Deist would have to believe that these algebras are destined to play an important role in the foundations of phusics, and, indeed, investigators into the theory of elementary particles have lately begun to show interest in the exceptional Lie groups. We present detailed models for all the non-exceptional and exceptional algebras, in a form convenient for calculation, in order to facilitate possible future computations involving these structures.

The methods introduced for the classification of the complex Lie algebras provide, following Cartan, Weyl, and Harish-Chandra, an exhaustive description of the linear representations of the semi-simple algebras, together with useful information on the automorphism groups of the semi-simple algebras, which are themselves semi-simple Lie groups. An additional analysis, in which we follow the elegant volume of Bourbaki, expresses a general algebra in terms of a semi-simple algebra and a solvable algebra, in effect reducing the general theory of complex Lie algebras to that of solvable Lie algebras. Unfortunately, the structure theory of solvable algebras, like most algebraic theories of structures which are neither commutative nor semi-simple, remains, in spite of some successful partial analyses, largely a mystery, perhaps even an impenetrable mystery.

The analysis of complex semi-simple algebras concluded, we pass (still following Cartan!) to study the corresponding real algebras. These also may be completely enumerated. The necessary investigation, however, is still complicated, even with a complete theory of the complex algebras in hand. Our treatment follows that of Gantmacher, and is considerably simpler than Cartan's original discussion. The representation theory of the real semi-simple algebras follows as a corollary of the complex theory. At its conclusion, we apply our analysis of real algebras to obtain a complete description of compact topological groups.

The authors would like to thank Edwin Goldfarb and Leon Greenberg for their painstaking critique which led to numerous improvements in the original manuscript.

Contents

Section O 1

A. Algebra 1
B. Functions of Operators 9
C. Analysis 12
D. Topology 12
E. Topological Groups 19

Part I. Lie Groups and Their Lie Algebras 20

1. Manifolds 20
2. Lie Groups: Local and Global Properties 37
3. The Lie Algebra 47
4. The Exponential Function 59
5. The Campbell-Baker-Hausdorff Theorem 65
6. Subgroups and Sub-algebras 72
7. Important examples—the classical matrix groups 81

Part II. Complex Semi-Simple Lie Algebras 86

1. Basics 86
2. Cartan's Criterion 93
3. Structure of Simple Complex Algebras 100
4. Models of the Simple Complex Algebras 122
5. Isomorphisms and Automorphisms 138
6. Representations of Complex Semi-Simple Algebras 142
7. The radical splitting theorem. Existence of Lie Groups with a given Lie algebra 159

Part III. Real Semi-Simple Lie Algebras 172

1. Structure and Representations of Simple Real Algebras 172
   A. Analysis of the conjugations in the Lie Algebra of type A_{n-1} 184
   B. Analysis of the conjugations in the Lie Algebra of types B_m, C_m and D_m 188
   A′. Analysis of the conjugations in the Lie Algebra of type G_2 201
   B′. Analysis of the conjugations in the Lie Algebra of type E_8 202
   C′. Analysis of the conjugations in the Lie Algebra of type E_7 205
   D′. Analysis of the conjugations in the Lie Algebra of type E_6 207
   E′. Analysis of the conjugations in the Lie Algebra of type F_4 213

2. Compact Connected Lie Groups 220

Index 227

Section 0

It is the purpose of this section to record, for later reference, some of the results from algebra, analysis, and topology which will be needed in the text. Proofs will be minimal or nonexistent.

A. Algebra

In what follows, V is an n-dimensional vector space (n finite) over the real or complex numbers.

1. Definition. A *bilinear form* on a vector space V (real or complex) is a mapping $[x, y] \to (x, y)$ of $V \times V$ into the set of (real or complex) numbers, such that

$$(ax + bx', y) = a(x, y) + b(x', y)$$

and

$$(x, ay + by') = a(x, y) + b(x, y')$$

for all vectors x, x', y, y' and scalars a, b.

2. Definition. A bilinear form is *symmetric* if

$$(x, y) = (y, x)$$

for all vectors x, y. A bilinear form is *skew-symmetric* if

$$(x, y) = -(y, x)$$

for all vectors x, y.

3. Definition. A *Hermitian form* on a complex vector space V is a mapping $[x, y] \to (x, y)$ of $V \times V$ into the set of complex numbers such that

$$(x, y) = \overline{(y, x)}$$

and

$$(ax + bx', y) = a(x, y) + b(x', y)$$

so that

$$(x, ay + by') = \bar{a}(x, y) + \bar{b}(x, y')$$

for all vectors x, x', y, y' and scalars a, b.

Remark. If (x, y) is Hermitian, then (x, x) is real for all vectors x. In coordinate form, (if x and y have coordinates x^i and y^i $[i = 1, \ldots, n]$ respectively) we have

$$(x, y) = \sum_{i,j} a_{ij}x^i\bar{y}^j$$

where $a_{ij} = \bar{a}_{ji}$. Conversely, such a matrix (a_{ij}) is called a Hermitian matrix and this equation defines a Hermitian form if (a_{ij}) is a Hermitian matrix. Definitions 1 and 2 can similarly all be put into coordinate form as follows:

(1) Bilinear	$(x, y) = \Sigma a_{ij}x^iy^j$	a_{ij}	(arbitrary matrix)
(2) Symmetric	$(x, y) = \Sigma a_{ij}x^iy^j$	$a_{ij} = a_{ji}$	(symmetric matrix)
(3) Skew-symmetric	$(x, y) = \Sigma a_{ij}x^iy^j$	$a_{ij} = -a_{ji}$	(skew-symmetric matrix)
(4) Hermitian	$(x, y) = \Sigma a_{ij}x^i\bar{y}^j$	$a_{ij} = \bar{a}_{ji}$	(Hermitian matrix).

4. Definition. A bilinear form (x, y) is called *non-singular* if, for every $x_0 \in V$ $(x_0 \neq 0)$, the linear form (x, x_0) is not identically 0 in x. Otherwise the bilinear form is said to be *singular*.

Remark. If we put $f_{x_0}(x) = (x, x_0)$, the mapping $x_0 \to f_{x_0}$ is clearly a linear map from V into the dual space V^* (of all linear functions on V). Definition 4 thus says that a bilinear form is non–singular if the map $x_0 \to f_{x_0}$ is $1 - 1$.

5. Lemma. Let (x, y) be a non-singular bilinear form on V. Let f be a linear function from V into the scalars of V. Then there is a unique $x_0 \in V$ such that

$$f(x) = (x, x_0).$$

Proof. This follows immediately from the above remark. For, $x_0 \to f_{x_0}$

is $1-1$ from V into V^*. Furthermore, V^* has the same dimension as V. Thus $x_0 \to f_{x_0}$ is onto. Q.E.D.

Remark. A bilinear form is non-singular if and only if, in coordinate form, (see Formula (1) above), its matrix (a_{ij}) is non-singular: $\det (a_{ij}) \neq 0$. This remark gives an easy proof of the statement that if a bilinear form (x, y) is non-singular, then for any $x_0 \in V (x_0 \neq 0)$, the linear form (x_0, x) is not identically 0 in x. Thus, we may interchange the roles of the first and second vector argument.

If V_1 is a subspace of V, then any bilinear form on V clearly induces a bilinear form on V_1 if we restrict the arguments to V_1. However, it is not true that the restriction of a non-singular form is non-singular. For example, if we define

$$(x, y) = x_1 y_1 - x_2 y_2 \qquad (\text{in } R^2)$$

and if V_1 is the space spanned by (1, 1), then the restriction of (x, y) to V_1 is identically 0. This observation leads to the following definition and lemma.

6. Definition. Let (x, y) be a non-singular bilinear form on V, and let V_1 be a subspace of V. Then V_1 is called a *non-singular* subspace (relative to the bilinear form) if the restriction of the form to V_1 is non-singular.

7. Lemma. Let (x, y) be a non-singular bilinear form on V, and let V_1 be a non-singular subspace of V. Let $V_1^\perp$ (the orthogonal complement of V_1) be $\{x | x \in V \text{ and } (x_0, x) = 0 \text{ for all } x_0 \in V_1\}$. Then V is the direct sum of V_1 and $V_1^\perp$. Furthermore, V_1 is also non-singular.

Proof. Using the remarks following Definition 4 and Lemma 5, we see that V_1 (assumed k-dimensional with $0 \leqslant k \leqslant n$) defines a k-dimensional subspace of the dual space of linear functions on V. Thus, $V_1^\perp$ is $(n-k)$-dimensional. Furthermore, if $x_0 \in V_1 \cap V_1^\perp$ it follows that $(x, x_0) = 0$ for all $x \in V_1$. Thus $x_0 = 0$ since V_1 is non-singular. Hence $V_1 \cap V_1^\perp = (0)$, and V is the direct sum of V_1 and $V_1^\perp$. If a basis $u_1, \ldots, u_k$ is chosen for V_1, and a basis $u_{k+1}, \ldots, u_n$ is chosen for $V_1^\perp$, then the matrix of (x, y) with respect to the basis $u_1, \ldots, u_n$ (see Formula (1)) is seen to have the form

$$\Gamma = \left(\begin{array}{c|c} A & 0 \\ \hline C & B \end{array} \right)$$

where Γ, A, and B are the matrices of (x, y) on V, V_1 and $V_1^\perp$ respectively. Since

$$0 \neq \det \Gamma = \det A \det B$$

we have $\det B \neq 0$. This shows that V_1 is non-singular by the remark following the proof of Lemma 5. Q.E.D.

Remark. It is clear that Definitions 4 and 6 apply to Hermitian forms, and that Lemmas 5 and 7 are true for Hermitian forms. Indeed, the proofs remain virtually unchanged.

8. Definition. Let (x, y) be either a symmetric form on a real vector space or a Hermitian form on a complex vector space. This form is said to be *positive definite* if $(x, x) > 0$ for all $x \neq 0$, $x \in V$.

Note that a positive definite form is non-singular. Note further that the restriction of a positive definite form to a subspace is clearly also positive definite. Thus, for a positive definite form, *all subspaces are non-singular*, and Lemma 7 applies. Furthermore, by symmetry, or Hermitian symmetry, it is clear that $V_1^{\perp\perp} = V_1$. For convenience, we restate these facts as a lemma.

9. Lemma. Let (x, y) be either a positive definite symmetric form on a real vector space V, or a positive definite Hermitian form on a complex vector space V. Then for any subspace $V_1 \subseteq V$ of dimension k, the set $V_1^{\perp} = \{x | (x_0, x) = 0 \text{ for all } x_0 \in V_1\}$ is $(n - k)$-dimensional, and V is the direct sum of V_1 and $V_1^{\perp}$. Furthermore, we have $V_1^{\perp\perp} = V_1$.

The existence of an "orthonormal" basis relative to any symmetric real form or to any Hermitian complex form is given by the following lemma.

10. Lemma. Let (x, y) be non-singular and either complex Hermitian, or real symmetric. Then there exists a basis $u_1, \ldots, u_n$ such that $(u_i, u_j) = 0$ $(i \neq j)$ and $(u_i, u_i) = \pm 1$.

Remark. In coordinate form, this lemma states that any such form, say,

$$Q(x, y) = \sum a_{ij} x_i \bar{y}_j$$

may be reduced to the standard form

$$Q'(x, y) = \sum_{i=1}^{k} x_i \bar{y}_i - \sum_{i=k+1}^{n} x_i \bar{y}_i$$

by applying a linear change of variables to x and the same change of variables to y.

Proof. We first note that (x, x) is not identically 0. (Left to the reader.) Choose an x_0 for which $Q(x_0, x_0) \neq 0$, and let $V_1 = (x_0)$ (the space spanned by x_0). Then V_1 is non-singular, and so is $V_1^{\perp}$ by Lemma 7. Now set

$u_1 = x_0/\sqrt{|(x_0, x_0)|}$, and use induction on the dimension of the underlying space V, applying the inductive step to the space $V_1^{\perp}$, to obtain the result. Q.E.D.

We now show that the number of basis elements u_i in any orthonormal basis with $(u_i, u_i) = 1$ is an invariant of the form.

11. Definition. Let (x, y) be either real symmetric, or complex Hermitian. The signature m of (x, y) is the maximum dimension of a subspace V_1 on which (x, y) is positive definite.

12. Lemma. Let (x, y) be either real symmetric or complex Hermitian, and suppose $u_1, \ldots, u_n$ is a basis of V such that $(u_i, u_j) = 0$ if $i \neq j$, $(u_i, u_i) = 1$ for $i = 1, \ldots, k$ and $(u_i, u_i) = -1$ for $i = k + 1, \ldots, n$. Let m be the signature of (x, y). Then $m = k$.

Proof. Clearly $k \leqslant m$. Let W_- denote the subspace of V spanned by $u_{k+1}, \ldots, u_n$. Then $(v, v) < 0$ for $v \in W_-$, $v \neq 0$. If W is a space of maximal dimension m on which (x, y) is positive definite, W and W_- have no non-zero vector in common. Hence $m + n - k \leqslant n$, i.e., $m \leqslant k$. Q.E.D.

13. Corollary. Suppose u is given, and $V_0 = \{v \in V | (v, u) = 0\}$. Then
a) if $(u, u) > 0$, the signature of (x, y) on V_0 is one less than the signature of (x, y) on V.
b) if $(u, u) < 0$, the signature of (x, y) on V_0 is equal to its signature on V.

We now review some of the basic definitions and results concerning operators and their eigenvalues. In what follows, all operators L are linear from V into V. If coordinates are chosen in V we may represent L by an $n \times n$ matrix A and obtain corresponding statements about matrices A.

14. Definition. Let the vector space be complex, i.e., admit complex scalars. If $Lx_0 = \lambda x_0$ with $x_0 \neq 0$, we say that x_0 is an *eigenvector* of L, and λ its corresponding *eigenvalue*. The set $\sigma(L)$ of eigenvalues of L is called the *spectrum* of L. If λ is an eigenvalue of L, then the subspace $V_\lambda = \{x | L(x) = \lambda x\}$ is called an *eigenspace* of L. Similarly, the space $M_\lambda = \{x | (L - \lambda I)^N x = 0 \text{ for sufficiently large } N\}$ is called the *generalized eigenspace* of L (corresponding to the eigenvalue λ).

Remark. The spectrum of L is the (finite) set of solutions of the n-th degree polynomial equation

$$\det(L - xI) = 0. \tag{1}$$

Had we considered real vector spaces, this same equation would have defined the eigenvalues of L, except that the real solutions x would have been sought. But the theory of polynomial equations demands *complex* numbers as the natural scalars for simplicity.

15. Definition. Let L be an operator on a complex vector space V. Then L is *diagonalizable* if there exists a basis $u_1, \dots, u_n$ of V consisting of eigenvectors:

$$L(u_i) = \lambda_i u_i .$$

Remark. The scalars λ_i are necessarily the solutions (with correct multiplicity) of the equation

$$\det (L - \lambda I) = 0 .$$

The simplest criterion that an operator be diagonalizable is given by the following well known result, whose proof we omit.

16. Lemma. Let the equation $\det (L - \lambda I) = 0$ have n distinct roots $\lambda_1, \dots, \lambda_n$. Then L is diagonalizable.

17. Definition. A *unitary* n-dimensional space V is a complex n-dimensional vector space V together with a positive definite Hermitian form (x, y) on V. (x, y) is often called the *inner product* of x and y.

Remark. It follows immediately from Lemmas 10 and 12 that basis vectors u_i may be chosen so that $(u_i, u_j) = 0$ $(i \neq j)$ and $(u_i, u_i) = 1$. If coordinates are taken with respect to such a basis, we have

$$(x, y) = \sum_i x^i \overline{y}^i .$$

For unitary spaces, we shall always take coordinates with respect to such orthonormal basis vectors unless otherwise stated.

18. Lemma. Let L be an operator on the unitary space V. Then there exists a unique operator M on V such that, for all x and y in V

$$(L(x), y) = (x, M(y)). \tag{1}$$

Proof. For fixed y, the function $N(x) = (L(x), y)$ is linear in x. By Lemma 5 and the remark preceding Definition 8 we have $N(x) = (x, x_0)$ for some unique x_0 (depending on y): $x_0 = M(y)$. Thus we have (1), and the uniqueness of $M(y)$. The proof that $M(y)$ is linear is left to the reader. Q.E.D.

19. Definition. Let L be an operator on the unitary space V. The *ad-*

joint of L, denoted L^* is the unique operator (given by Lemma 18) which satisfies the equation

$$(L(x), y) = (x, L^*(y)). \tag{2}$$

Remark. In coordinates, the matrix of L^* is simply the conjugate transpose of the matrix of L. (This is true only when orthonormal basis vectors are used.) We write, for matrices, $A^* = \bar{A}^t$.

20. Lemma. Let L and M be operators on the unitary space V. Then

$$L^{**} = L,\ (aL)^* = \bar{a}L^*,\ (L + M)^* = L^* + M^*,\ (LM)^* = M^*L^*.$$

Proof. Left to the reader.

Remark. L^* is the analogue, to operators, of the complex conjugate $\bar{z}$ of a complex number.

21. Definition. An operator L on a unitary space V is called *Hermitian*, if $L^* = L$. Equivalently, L is Hermitian if $(L(x), y) = (x, L(y))$ for all $x, y \in V$.

The following lemma relates Hermitian forms and Hermitian operators (cf. Definition 3).

22. Lemma. Let L be a Hermitian operator on the unitary space V. Then the form $\{x, y\} = (L(x), y)$ is a Hermitian form. Conversely, if $\{x, y\}$ is a Hermitian form, there exists a unique Hermitian operator L on V, such that $\{x, y\} = (L(x), y)$.

We omit the easy proof.

The following important lemma contains useful information on the eigenvalues and eigenvectors of Hermitian operators.

23. Lemma. Let L be a Hermitian operator on the unitary space V. Then

(a) The eigenvalues of L are all real.
(b) The distinct eigenspaces of L are orthogonal.
(c) L is diagonalizable.
(d) V is the direct sum of the eigenspaces of L.

Proof.

(a) If $L(x) = \lambda x$, with $x \neq 0$, then $(L(x), x) = (x, L(x))$ or $\lambda(x, x) = \bar{\lambda}(x, x)$.

Hence λ is real.

(b) If $\lambda_1 \neq \lambda_2$, and $L(x) = \lambda_1 x$, $L(y) = \lambda_2 y$, then

$$\lambda_1(x, y) = (L(x), y) = (x, L(y)) = \bar{\lambda}_2(x, y) = \lambda_2(x, y).$$

Hence $(x, y) = 0$.

(c) We work by induction on the dimension n of V. The case $n = 1$ is trivial. If λ_1 is an eigenvalue of L with eigenvector u_1, let $V_0 = \{x|(x, u_1) = 0\}$. Then L maps V_0 into V_0, since if $x \in V_0$, $(L(x), u_1) = (x, L(u_1)) = \bar{\lambda}_1(x, u_1) = 0$. Clearly L is Hermitian on V_0 since $(L(x), y) = (x, L(y))$ holds in all V, hence in V_0. Therefore, by induction, there is a basis $u_2, \ldots, u_n$ of V_0 for which $L(u_i) = \lambda_i u_i$ $(i = 2, \ldots, n)$. Thus, $u_1, \ldots, u_n$ are basis vectors of V for which $f(u_i) = \lambda_i u_i$. This completes the proof.

(d) This follows immediately from (c). Q.E.D.

For the analogy suggested by the remark preceding Definition 21, Hermitian operators are the analogues of real numbers. We now proceed to discuss positive definite Hermitian forms—the analogues of positive real numbers.

24. Definition. Let L be a Hermitian operator on the unitary space V. Then L is called *positive definite* if the Hermitian form $(L(x), y)$ is positive definite.

25. Lemma. Let L be a Hermitian operator on the unitary space V. Then L is positive definite if and only if all of the eigenvalues of L are positive.

This is clear from Lemma 23.

26. Lemma. Let L be a positive definite Hermitian operator on the unitary space V. Then there exists a unique "square root" M of L satisfying

(a) $M^2 = L$
(b) M is a positive definite Hermitian operator.
Further, this mapping M satisfies the condition:
(c) If L commutes with an operator N on V, then M also commutes with N.

Proof. If the matrix of L with respect to a suitable orthonormal basis is

$$(1) \qquad [L] = \begin{pmatrix} \lambda_1 & & & 0 \\ & \cdot & & \\ & & \cdot & \\ 0 & & & \lambda_n \end{pmatrix}$$

with $\lambda_i > 0$, we may define M as the operator whose matrix is:

$$[M] = \sqrt{[L]} = \begin{pmatrix} \sqrt{\lambda_1} & & & 0 \\ & \cdot & & \\ & & \cdot & \\ 0 & & & \sqrt{\lambda_n} \end{pmatrix} \tag{2}$$

Then $M^2 = L$ and M is a positive definite Hermitian operator. The uniqueness of M may readily be seen by applying Lemma 23 to M, squaring M and comparing with (1). The proof of (c) involves a matrix calculation: $[L]\,[N] = [N]\,[L]$ $([N] = (n_{ij}))$, holds if and only if for any i and j $(1 \leqslant i, j \leqslant n)$, we have $n_{ij} = 0$ or $\lambda_i = \lambda_j$. While $\sqrt{[L]}\,[N] = [N]\sqrt{[L]}$ hold if and only if for any i and j we have $n_{ij} = 0$ or $\sqrt{\lambda_i} = \sqrt{\lambda_j}$. Since positive square roots are taken, we have the result. Q.E.D.

B. Functions of Operators

We take for granted some of the usual extensions to a matrix or operator variables of certain elementary topological and analytic notions. Since operators act on finite dimensional spaces, and matrices have finitely many (n^2) entries, such notions as, say, a continuous matrix-valued mapping $M(t)$ of a real (or complex) variable, an analytic operator mapping of a complex variable, an integral $\int_a^b M(t)\,dt$, etc., are easily extended by a component-by-component definition. In particular, we note that we may readily define a norm of an operator $|A|$ satisfying the usual rules:

$$|A| > 0 \quad \text{if} \quad A \neq 0; \qquad |0| = 0; \quad |A + B| \leqslant |A| + |B|;$$

$$|cA| = |c|\,|A|, \ |AB| \leqq |A|\,|B|.$$

For example, if $A = (a_{ij})$ is an $n \times n$ matrix, we may define

$$|A| = |(a_{ij})| = n \max_{i,j} |a_{ij}|$$

We now consider how a small perturbation of an operator effects its eigenvalues.

27. Lemma. Let the distinct eigenvalues of the operator A be $\lambda_1, \ldots, \lambda_k$. Suppose that $\varepsilon_i > 0$, and $C_i = \{\lambda | |\lambda - \lambda_i| < \varepsilon_i\}$ $(i = 1, \ldots, k)$, and that the C_i are disjoint. Then there exists $\delta > 0$ such that if B is any operator with $|B - A| < \delta$, then each eigenvalue of B is in one of the circles C_i, and each circle C_i contains at least one eigenvalue of B. Furthermore, suppose that λ is one of the λ_i, and C the corresponding C_i, and that $\mu_1, \ldots, \mu_r$ are the eigenvalues of the operator B in the circle C. Let $M(\lambda)$ and $M(\mu_i)$ be the generalized eigenspaces of A and B respectively (cf. Definition 14). Then, dimension $M(\lambda) = \sum_i$ dimension $M(\mu_i)$. In particular, if B has no more eigenvalues than A, then each circle C_i contains precisely one eigenvalue μ_i of B, and the spaces $M(\lambda_i)$ and $M(\mu_i)$ have the same dimension.

If the operator B varies *analytically* with z, we can give a more precise statement. No proof is given.

28. Lemma. Let $A(z)$ be an analytic operator-valued function of z defined in some neighborhood of $z = 0$. Let λ_0 be an eigenvalue of $A_0 = A(0)$, and let $\varepsilon > 0$ (possibly infinite) be the distance from λ_0 to the nearest other eigenvalue of A_0. Then there exists a fractional power series $\lambda(z^{1/k})$, with $\lambda(0) = \lambda_0$, such that

(1) for all z in a sufficiently small neighborhood of $z = 0$, all of the k values of $\lambda(z^{1/k})$ are eigenvalues of $A(z)$, and

(2) for all z in a sufficiently small neighborhood of 0, if μ is an eigenvalue of $A(z)$ with $|\mu - \lambda_0| < \frac{\varepsilon}{2}$, then μ is one of the values of $\lambda(z^{1/k})$.

As an immediate consequence of Lemma 28, we have the following lemma concerning the number of distinct eigenvalues of $A(z)$.

29. Lemma. Let $A(z)$ be an operator-valued analytic function of z defined in some open set, and let $m(z)$ be the number of distinct eigenvalues of $A(z)$. Then $m(z)$ is constant $(= c)$ except at an isolated set S on which $m(z) < c$.

We now consider analytic operator-valued functions of an operator variable. The definition we use follows.

30. Definition. Let $w = f(z)$ be an analytic function defined in the domain D bounded by a finite collection C of Jordan curves. If A is any operator all of whose eigenvalues λ are in D, we define

$$W = f(A) = \frac{1}{2\pi i} \int_C f(z)\,(zI - A)^{-1}\,dz.$$

Remark. Thus, the domain of definition of the analytic operator-valued function $f(A)$ is the set of all operators A, all of whose eigenvalues are in the domain of definition D of $f(z)$; i.e., all operators whose spectrum is contained in D.

The relationship of this definition to power series is given by the following lemma, which we do not prove here.

33. Lemma. Let $f(z)$ be analytic with $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ valid in the circle C_R: $\{z \mid |z - z_0| < R\}$. Suppose A is an operator whose spectrum is in C_R. Then

$$f(A) = \sum_{n=0}^{\infty} a_n(A - z_0 I)^n.$$

In particular, the usual power series expressions for the familiar analytic functions are valid for operators. For example

$$e^A = 1 + A + \frac{1}{2!} A^2 + \cdots$$

The next lemmas (also unproven here) show that analytic identities remain valid even for operator values of the argument.

34. Lemma. Let $f(z)$ be defined and analytic in the open set D, and let $g(z)$ be defined and analytic in the open set E. Suppose the $g(E) \subseteq D$, so that $h(z) = f(g(z))$ is defined and analytic in E. Suppose that the eigenvalues of A are $\lambda_1, \ldots, \lambda_n$ (counting multiplicities) and that $\lambda_i \in E$ $(i = 1, \ldots, n)$. Then the eigenvalues of $g(A)$ are $g(\lambda_1), \ldots, g(\lambda_n)$ counting multiplicities. In particular $f(g(A))$ is defined whenever $g(A)$ is.

35. Lemma. Let $f(z)$, $g(z)$ be given as in Lemma 34, and let the spectrum of A be in E. Then

$$h(A) = f(g(A)).$$

Remark. The preceding lemma states that functional identities are preserved for operator-valued functions. For example, taking $f(z) = z^2$, $g(z) = e^z$, so that $h(z) = f(g(z)) = e^{2z}$, we have

$$(e^A)^2 = e^{2A}.$$

Similarly, if $\log z$ is defined for $|z - 1| < 1$ by the usual series, we have

$$e^{\log A} = A; \quad \log e^A = A$$

for the appropriate set of operators A.

As a final non-surprising result, we state the following lemma, also not proven here.

36. Lemma. Let $f(z, w)$ be analytic for $z \in D$, $w \in E$, and let $h(z, w) = \frac{\partial}{\partial z} f(z, w)$. Suppose the spectrum of A is in E, and define

$$M(z) = f(z, A).$$

Then $\frac{d}{dz} M(z) = h(z, A)$.

C. Analysis

For later convenience, we state the implicit function theorem for C^∞ functions.

37. Lemma. Let $f^1(x^1, \ldots, x^n), \ldots, f^n(x^1, \ldots, x^n)$ be n C^∞ functions of n variables $x^1, \ldots, x^n$, defined in a neighborhood of $x_0^1, \ldots, x_0^n$. Let

$$J(x^1, \ldots, x^n) = \det \left(\frac{\partial f^i}{\partial x^j} \right)$$

be the Jacobian determinant, and suppose that $J(x_0^1, \ldots, x_0^n) \neq 0$. Let $w_0^1 = f_1(x_0^1, \ldots, x_0^n), \ldots, w_0^n = f_n(x_0^1, \ldots, x_0^n)$. Then there is a neighborhood V of $(x_0^1, \ldots, x_0^n)$ and there are n C^∞ functions $g^1(w^1, \ldots, w^n), \ldots, g^n(w^1, \ldots, w^n)$ satisfying $g^1(w_0^1, \ldots, w_0^n) = x_0^1, \ldots, g^n(w_0^1, \ldots, w_0^n) = x_0^n$ defined in a neighborhood of $w_0^1, \ldots, w_0^n$ such that for any $(w^1, \ldots, w^n)$ in this neighborhood, the equations

$$w^1 = f_1(x^1, \ldots, x^n), \ldots, w^n = f_n(x^1, \ldots, x^n)$$

have the unique solution $x^1 = g^1(w^1, \ldots, w^n), \ldots, x^n = g^n(w^1, \ldots, w^n)$ such that $(x^1, \ldots, x^n) \in V$.

D. Topology

We shall now summarize some of the definitions and results concerning the fundamental group and the covering spaces of a topological space. Proofs will be either sketched or omitted. In what follows X is a connected topological space and P_0 is a given point in X, called the base point. We let I designate the closed unit interval [0, 1].

38. Definition. A *closed curve* f is a mapping $f: I \to X$ such that $f(0) = f(1) = P_0$ (= base point). A *curve* is simply a mapping $f: I \to X$.

39. Definition. Let f and g be closed curves in X. We say that f is homotopic to $g (f \cong g)$ if there exists a one-parameter family of curves f_s $(0 \leqslant s \leqslant 1)$ such that

(1) $f_s(t)$ is a continuous mapping, regarded as a map of $I \times I$ into X.
(2) $f_0(t) = f(t); f_1(t) = g(t); f_s(0) = f_s(1) = P_0$.

We shall occasionally employ the notion of homotopy for curves which are not necessarily closed. In that case we write $f \cong g$ if $f(0) = g(0); g(1) = g(1)$ and if there exists a family f_s of curves satisfying (1) and

(2)′ $\quad f_0(t) = f(t); \quad f_1(t) = g(t); \quad f_s(0) = f(0); f_s(1) = f(1).$

In this case we say we have a homotopy with fixed end-points.

40. Lemma. The relation $f \cong g$ is an equivalence relation.

41. Notation. The equivalence class of homotopic curves containing the closed curve f is designated $[f]$. If f_0 is the constant closed curve ($f_0(t) = P_0$ for all t), we put $[f_0] = [P_0] = e$.

42. Definition. Let X be a path-wise connected space. (I.e., if $P_1, P_2 \in X$, there exists a mapping $f: I \to X$ such that $f(0) = P_1, f(1) = P_2$.) Then X is said to be *simply connected* if all closed curves are homotopic. Equivalently, X is simply connected if $[f] = e$ for all closed curves f.

Remark. This definition is easily shown to be independent of the base point P_0.

43. Definition. A space X is *locally simply connected* if for any neighborhood U of X and any $P \in U$, there exists a simply connected neighborhood V such that $P \in V \subseteq U$.

For the remainder of this section, we shall consider only *spaces that are path-wise connected, locally simply connected, and Hausdorff spaces.*

44. Definition. If f and g are closed curves in X, the curve $f * g$ is the closed curve defined by

$$(f * g)(t) = f(2t), \qquad 0 \leqslant t \leqslant \tfrac{1}{2}$$

$$(f * g)(t) = g(2t - 1), \qquad \tfrac{1}{2} \leqslant t \leqslant 1.$$

The inverse closed curve f^{-1} is defined by the formula

$$f^{-1}(t) = f(1 - t), \qquad 0 \leqslant t \leqslant 1.$$

Remark. $f * g$ is merely "one curve followed by the other". f^{-1} reverses the orientation of f.

Even if f and g are not closed curves, $f * g$ is defined as a continuous curve provided $f(1) = g(0)$. f^{-1} is clearly defined for all, not necessarily closed, curves.

45. Lemma. if $f_1 \cong f_2$, $g_1 \cong g_2$, then $f_1 * g_1 \cong f_2 * g_2$ and $f_1^{-1} \cong f_1^{-1}$. Hence we may define $[f_1] * [g_1] = [f_1 * g_1]$ and $[f_1]^{-1} = [f_1^{-1}]$. Under this definition of multiplication and inverse, the set of equivalence classes of closed curves forms a group with identity $e = [P_0]$.

46. Definition. The *fundamental group* $\pi_1(X)$ is the group of equivalence classes of closed curves defined in Lemma 45. If the base point P_0 is to be made explicit, we write $\pi_1(X, P_0)$.

Remark. If P_0 and P_1 are arbitrary in X, then $\pi_1(X, P_0)$ is isomorphic to $\pi_1(X, P_1)$. Thus, $\pi_1(X)$ is well defined, as an abstract group. However, there is no unique natural homomorphism of $\pi_1(X, P_0)$ into $\pi_1(X, P_1)$.

47. Lemma. Let X and Y be spaces with base points P_0 and Q_0 respectively, and let $p: X \to Y$ with $p(P_0) = Q_0$. Then

(1) If f is a closed curve in X, pf is a closed curve in Y.
(2) If $f \cong g$ then $pf \cong pg$.
(3) $p(f * g) = (pf) * (pg)$.

Thus, p induces a map of $\pi_1(X, P_0)$ into $\pi_1(Y, Q_0)$ according to the correspondence $[f] \to [pf]$. This map is homomorphism.

48. Definition. If $p: X \to Y$ is given as in Lemma 47, we define the homomorphism $p_*: \pi_1(X, P_0) \to \pi_1(Y, Q_0)$ by the formula $p_*[f] = [pf]$.

49. Lemma. If $p: X \to Y, q: Y \to Z$ with $p(P_0) = Q_0, q(Q_0) = R_0$, and if P_0, Q_0, and R_0 are base points in X, Y, and Z respectively, then

$$(qp)_* = q_* p_*; \qquad i_* = i$$

where i is taken as a generic symbol for the identity map.

We shall now introduce the notion of a covering of a space, and show how these are related to the fundamental group.

50. Definition. Let $p: Y \to X$. Then Y is said to *cover* X with the *projection* p if for any $P \in X$, there exists a simply connected neighborhood U of P such that $p^{-1}(U)$ is non-empty, and such that p is a homeomorphism of each connected component of $p^{-1}(U)$ onto U.

Remark. It follows that $p^{-1}(P)$ is non-empty and discrete. It is clear from the definition that if $p: Y \to X$ is a covering, then for each $P \in X$, the set $p^{-1}(P)$ has cardinality n which is independent of P. Indeed, this cardinality is locally constant, and X is connected. If $n = 1$ then p is clearly a homeomorphism.

The easiest and most well known non-trivial covering is given by $p:R \to S$, where R = real numbers, S = unit circle ($|z| = 1$), and $p(x) = e^{2\pi ix}$.

Base points P_0 and Q_0 in X and Y are always chosen so that $p(Q_0) = P_0$. We shall use the expression "Q is over P" when $p(Q) = P$. Similarly, we "lift P to Q", etc.

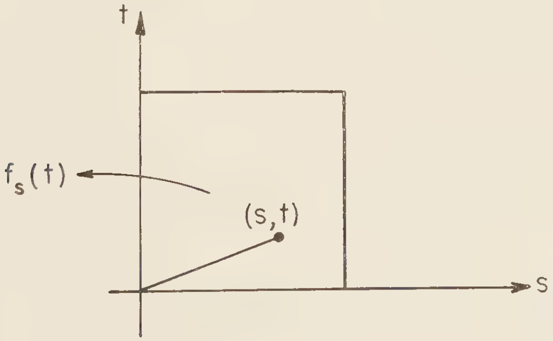
51. Lemma. Let $p:Y \to X$ be a covering and let $f:I \to X$ be a (not necessarily closed) path in X. Let $P_0 = f(0)$ and let Q_0 be over P_0. Then there exists a unique curve $g:I \to Y$ such that $pg = f$ and $g(0) = Q_0$. Briefly a curve in X may be lifted uniquely with an arbitrary starting point lying over P_0.

Proof. Existence: Cover $f(I)$ (the curve as point set) with neighborhoods U as in Definition 50. By compactness, we may cover with finitely many such U. Since p is a local homeomorphism, we lift f step by step and thus lift it in the large.

Uniqueness: Left to the reader.

52. Lemma. Let $p:Y \to X$ be a covering and let f_1 and f_2 be curves in X with $f_1 \simeq f_2$. Let g_1 and g_2 be lifts of f_1 and f_2 respectively, with $g_1(0) = g_2(0)$. Then g_1 and g_2 are homotopic (with end-points fixed). Thus there is a one parameter mapping $H_s:I \to Y$ such that $H_s(t)$ is continuous in s and t, $H_0(t) = g_1(t)$; $H_1(t) = g_2(t)$, $H_s(0) = g_1(0)$ and $H_s(1)$ = constant. In particular, $g_1(1) = g_2(1)$ $(= H_s(1))$.

Proof. Suppose $f_s(t)$ is the homotopy connecting $f_1(t)$ and $f_2(t)$. We shall lift $f_s(t)$ to construct $H_s(t)$. This may be done as follows. For any (s, t), the



line from (0, 0) to (s, t) gives a curve in X. Lift it, with starting point $g_1(0)$, and let $H_s(t)$ be the endpoint of this lifted curve. This yields a transformation $H_s(t)$ of $I \times I$ into Y. To show that $H_s(t)$ is a continuous map of $I \times I$ into Y, we cover the curve in X which corresponds to the line to (s, t) with neighborhoods as in the proof of Lemma 51. We then cover the line from (0, 0) to (s, t) with neighborhoods whose image (under $f_s(t)$) are wholly within these neighborhoods. It is clear that if (s, t) is changed slightly to (s', t'), then the line to (s', t') is still wholly within these neighborhoods. The proof of Lemma 51 then shows that $H_s(t)$ is a continuous mapping. Finally, $H_s(1)$ is a lift of the constant curve $f_s(1)$, hence is a constant by uniqueness in Lemma 51. Q.E.D.

53. Corollary. $\pi_1(X)$ acts transitively on $p^{-1}(P_0)$ in a natural way.

Indeed, if $[f] \in \pi_1(X)$ and if $Q \in p^{-1}(P_0)$, then $(Q)[f]$ may be defined to be the unique $Q' \in p^{-1}(P_0)$ such that there exists a curve g over f with $g(0) = Q$ and $g(1) = Q'$. It is then clear that $((Q)[f_1])[f_2] = (Q)[f_1 * f_2]$ and $(Q)e = Q$. Finally, if Q_1 and Q_2 are over P_0, they may be joined by a curve $g(t)$ in Y. $g(t)$ is certainly the lift of $f(t) = pg(t)$. Hence $(Q_1)[f] = Q_2$. (Note that the definition of $f * g$ is such that we must make $\pi_1(X)$ act on $p^{-1}(P_0)$ on the right.)

54. Corollary. If X is simply connected, and $p: Y \to X$ is a covering, then p is a homeomorphism. Briefly, a simply connected space can only be covered, up to homeomorphism, by itself.

Proof. If $P_0 \in X$, $\pi(X)$ acts transitively on $p^{-1}(P_0)$. But $\pi(X)$ is trivial. Hence $p^{-1}(P_0)$ contains just one point. Hence p is a homeomorphism by the remark following Definition 50.

We shall now generalize Corollary 54 and show that the coverings of X are (up to homeomorphism) in one-one correspondence with the subgroups of $\pi(X)$.

55. Lemma. Let $p: Y \to X$ be a covering. Then $p_* : \pi_1(Y) \to \pi_1(X)$ is 1–1 into.

Proof. Suppose Q_0 is the base point of Y, and $p(Q_0) = P_0$ is the base point of X. If $p_*[g_1] = p_*[g_2]$ then $[pg_1] = [pg_2]$ or $pg_1 \cong pg_2$. By Corollary 52, the lifts of pg_1 and pg_2 are homotopic: $g_1 \cong g_2$. Hence, $[g_1] = [g_2]$ which gives the result. Q.E.D.

56. Corollary. Let $p: Y \to X$ be a covering. Then $\pi_1(Y)$ is isomorphic to a subgroup H of $\pi_1(X)$. H consists of precisely those elements α of $\pi_1(X)$

which keep Q_0 (the base point of Y) fixed: $Q_0\alpha = Q_0$. We also say that H is the homotopy group of Y.

57. Lemma. Let $p:Y_1 \rightarrow Y_2$ be a covering and let $p_2:Y_2 \rightarrow X$ be a covering. Then $p_1 = p_2p:Y \rightarrow X$ is a covering. Let the group of $p_i:Y_i \rightarrow X$ be G_i, where $G_i \subseteq \pi_1(X)$, as in Corollary 56. Then $G_1 \subseteq G_2$.

$$\begin{tikzcd} Y_1 \arrow[r, "p"] \arrow[rr, bend right, "p_1"'] \arrow[rr, dash, bend left=60, "G_1"] & Y_2 \arrow[r, "p_2"] \arrow[r, dash, bend left, "G_2"] & X \end{tikzcd}$$

Proof. It is clear from Definition 50 that $p_1:Y_1 \rightarrow X$ is a covering. Let f be any closed curve in X, starting at P_0. Lift it to f_2 in Y_2, starting at P_2, the base point of Y_2. Lift f_2 to f_1 in Y_1 starting at P_1, the base point of Y. Then f_1 is the lift of f to Y_1. Now suppose $[f] \in G_1$. This means (by Corollary 56) that f_1 is a closed curve in Y_1 (i.e., $(P_1)[f] = P_1$). Hence $pf_1 = f_2$ is closed in Y_2, and therefore $(P_2)[f] = P_2$. Hence $[f] \in G_2$. Q.E.D.

58. Lemma. Conversely, suppose Y_1 and Y_2 cover X, with projection p_1 and p_2 respectively, and subgroups G_1 and G_2 of $\pi_1(X)$ respectively. Suppose further that $G_1 \subseteq G_2$. Then there exists a unique map $p:Y_1 \rightarrow Y_2$ such that $p_1 = p_2p$ and such that Y_1 covers Y_2 with projection p.

Proof. Let $P \in Y_1$, and let f_1 be any curve in Y_1 joining the base point P_1 to P. Then $f = p_1f_1$ is a curve in X joining the base point P_0 to p_1P. Now lift this curve to Y_2 and we arrive at a curve joining the base point P_2 to a point over p_1P, and under P, which we define to be $p(P)$. It is an easy matter to show that $p(P)$ is defined independently of the curve f_1 joining P_1 to P. (Here is where the hypothesis $G_1 \subseteq G_2$ is used.) Hence $p(P)$ is well-defined. It is also clearly continuous, and it satisfies the equation $p_1 = p_2p$. What is required, then, is to prove that $p:Y_1 \rightarrow Y_2$ is a covering. This is simple, and we leave it to the reader. Q.E.D.

59. Lemma. Let $p_1:Y_1 \rightarrow X$ and $p_2:Y_2 \rightarrow X$ be coverings of X with the same subgroup G of $\pi_1(X)$. Then Y_1 and Y_2 are homeomorphic by a homeomorphism $p:Y_1 \rightarrow Y_2$ which satisfies $p_2p = p_1$.

Proof. By Lemma 58, there is a mapping $p:Y_1 \rightarrow Y_2$ such that $p_1 = p_2p$. We shall show that p is a homeomorphism. To do this, we show that there is only one point about P_2, the base point of Y_2. For, suppose P_1 (the base point of Y_1) and Q_1 are above P_2. Join them by a curve f_1 in Y_1, starting at P_1 and ending at Q_1. Then pf_1 is a closed curve in Y_2 and $p_1f_1 = p_2pf_1$ is a closed curve in X. Clearly, the lift to Y_2 of p_1f_1 is pf_1 and therefore this lift is closed. Hence $[p_1f_1]$ is in the subgroup G of Y_2. By hypothesis, it is in the group of Y_1, and therefore its lift to Y_1 is closed. But this says that f_1 is a closed curve and therefore $P_1 = Q_1$. Q.E.D.

The preceding results concern uniqueness of covering spaces. We now prove their existence.

60. Lemma. Let X be a space and let G be a subgroup of $\pi_1(X)$. Then there exists a covering $p:Y \to X$ with group G. In particular, taking $G = (e)$, it follows that there exists a simply connected covering of X.

Proof. If f_1 and f_2 are curves in X which start at the base point P_0 and have a common end-point, write $f_1 \sim f_2$ if and only if $[f_1f_2^{-1}] \in G$. (For the simply connected covering, this is simply $f_1 \cong f_2$.) Write $\{f_1\}$ for the equivalence class of f_1 under the relation $\sim$. We define $Y =$ the set of all equivalence classes $\{f\}$. We define $p:Y \to X$ by the formula $p\{f\} = f(1)$. If $\{f\}$ is any point in Y, we define a neighborhood of $\{f\}$ as follows: Choose any simply connected neighborhood U of $f(1)$. Let g be any curve with $g(0) = f(1)$, $g(t) \in U$. Then the U-neighborhood of $\{f\}$ consists of all elements $\{f * g\}$. It is any easy matter to show that this makes Y into a connected, locally simply connected Hausdorff space. (X is assumed to be one.) Further, $p:Y \to X$ is a covering and its fundamental group may be verified to be G. Q.E.D.

We now summarize the preceding lemmas and put them in a more revealing perspective.

61. Definition. Let $p_i:Y_i \to X$ $(i = 1, 2)$ be coverings. We say that these coverings are equivalent if there exists a homeomorphism $p:Y_1 \to Y_2$, which sends the base point of Y_1 into the base point of Y_2, such that $p_2p = p_1$.

62. Definition. Let $p_i:Y_i \to X$ $(i = 1, 2)$ be coverings, and suppose there is a covering $p:Y_1 \to Y_2$ (sending the base point of Y_1 into the base point of Y_2) such that $p_2p = p_1$. Then we write $Y_1 \gg Y_2$.

63. Lemma. Let X be a topological space, with fundamental group $\pi_1(X)$. Then there is a one to one correspondence between subgroups G of $\pi_1(X)$ and covering spaces (up to equivalence) of X. The covering $p:Y$

$\to X$ corresponds to the subgroup G consisting of all $[f] \in \pi_1(X)$ whose lift is a closed curve, and G is isomorphic to $\pi_1(Y)$. If $G_1 \leftrightarrow Y_1$, $G_2 \leftrightarrow Y_2$, then $G_1 \subseteq G_2$ if and only if $Y_1 \gg Y_2$. In particular, taking $G = (e)$, it follows that there exists a unique (up to equivalence) simply connected covering $\tilde{X}$, which covers all coverings of X. (I.e., $\tilde{X} \gg Y$ for all coverings Y). This space $\tilde{X}$ is called the *universal covering space* of X.

E. Topological Groups

In the last section of this book, we shall have occasion to use the existence of an invariant integral on a compact group. We now state the relevant definition and results.

64. Definition. A topological group G is a topological space, which is also a group under $\otimes$, and such that the mapping $(g, h) \to g \otimes h$ is continuous from $G \times G$ onto G, and the mapping $g \to g^{-1}$ is continuous from G onto G.

65. Lemma. Let G be a compact topological group satisfying the second axiom of countability. Then for every real-valued continuous function $F(g)$ defined on G, there is a real number $\int F(g)\, dg$ satisfying the following conditions:

(1) $\int [aF(g) + bG(g)]\, dg = a \int F(g)\, dg + b \int G(g)\, dg$.
(2) If $G(g) \geqslant 0$, but $F(g)$ is not identically 0, then $\int F(g)\, dg > 0$.
(3) $\int 1\, dg = 1$.
(4) If a and b are fixed group elements, $\int F(agb)\, dg = \int F(g)\, dg$.
(5) $\int F(g^{-1})\, dg = \int F(g)\, dg$.

Furthermore, there is only one such functional $F(g) \to \int F(g)\, dg$ which satisfies these conditions.

PART I

Lie Groups and Their Lie Algebras

1. Manifolds

All of the classical infinitesimal and local analysis involving functions of several variables may be generalized to analysis on a *manifold.* Since, by definition, a manifold is a locally Euclidean space this is almost a redundant statement.

1. Definition. An *n-dimensional manifold M* is a Hausdorff space, together with an open covering U_α of M and mappings $p_\alpha : U_\alpha \to R^n$, where we require that p_α is a homeomorphism of U_α onto the open subset $p_\alpha(U_\alpha)$ of R^n. The pair (p_α, U_α) is called a *coordinate system*, p_α is called the coordinate mapping, and if $P \in U_\alpha$, $x = p_\alpha(P)$ is called the *coordinate* of P in the coordinate system (p_α, U_α).

Thus, $x = (x^1, \ldots, x^n) = p_\alpha(P)$ is the (vector) coordinate of P, while $p_\alpha^{-1}(x)$ is the point P of M which has the coordinate x. By means of p_α and p_α^{-1} we are able to go back and forth from M to Euclidean space and are thereby able to define all the standard notions of local analysis on M by using the definition of these notions in Euclidean space. But, if our manifold is to support calculus and not merely topology, we must impose the following additional condition on its coordinate mappings.

2. Definition. A C^∞ manifold M is a manifold with the added restriction that $p_\beta p_\alpha^{-1}$ is C^∞ where it is defined (i.e. on the open set $p_\alpha(U_\alpha \cap U_\beta)$ if $U_\alpha \cap U_\beta \neq \phi$).

In more familiar terminology, we require that the transformation of coordinates connecting two overlapping coordinate systems be given by C^∞ functions. In the remainder of this book, *all manifolds are understood to be* C^∞. *In the remainder of the present section, M is understood to be a manifold of dimension n, and M′ a manifold of dimension n′.*

Remark. R^n is a (C^∞) manifold: take $U = U_1 = R^n$ and let $p = p_1$ be the identity mapping. So is any open subset G of R^n. Somewhat more generally, if V is any n-dimensional real vector space then V is a (C^∞) manifold. For if $e_1, \ldots, e_n$ is any basis, we have a (p, U) where U is V and $p(\lambda^1 e_1 + \cdots + \lambda^n e_n) = (\lambda^1, \ldots, \lambda^n)$. The transformation between these coordinates and the coordinates given by a different choice of basis is linear, hence C^∞.

The generalization of the notion of a C^∞ function or of a C^∞ mapping to manifolds is now clear—we express functions in terms of coordinates.

3. Definition. Let M, M' be manifolds of dimensions n and n' respectively and let f be a mapping of M into M'. Then we say that f is C^∞ at $P_0 \in M$ if for any coordinate system (p_α, U_α) of M and (p'_β, V_β) of M' such that $P_0 \in U_\alpha$, $P'_0 = f(P_0) \in V_\beta$, and $p_\alpha(P_0) = x_0$, the function $\tilde{f}(x) = p'_\beta f p_\alpha^{-1}(x)$ is C^∞ in some neighborhood of x_0. We say that f is a C^∞ map if it is C^∞ at each point of M.

In particular, putting $M' = R^{n'}$, we obtain the definition of a "C^∞ vector-valued function on M". Then, putting $n' = 1$ (resp. $n' = 2$) we obtain the definition of a "C^∞ real-valued (resp. complex-valued) function on M".

Briefly, a map f is C^∞ if, expressed in coordinates, it is described by a set of C^∞ functions. We note that f is C^∞ at P_0 if $\tilde{f}(x) = p'_\beta f p_\alpha^{-1}(x)$ is C^∞ at x_0 for even one pair of coordinate systems (p_α, U_α), (p'_β, V_β) with $P_0 \in U_\alpha$, $f(P_0) \in V_\beta$, $p_\alpha(P_0) = x_0$. This is clear from the formula

$$p'_\delta f p_\gamma^{-1} = (p'_\delta p'^{-1}_\beta)(p'_\beta f p_\alpha^{-1})(p_\alpha p_\gamma^{-1})$$

Thus, since the change of coordinates is C^∞, we see that if the expression for f in one pair of coordinate systems is C^∞, so is the expression for f in any other pair. In the remainder of this book, *all mappings are understood to be* C^∞, unless the contrary is explicitly specified.

Notation. If f is a mapping of M into M', we write $\tilde{f} = p'_\beta f p_\alpha^{-1}$ as in the previous definition. The mapping $\tilde{f}$ clearly depends on the choice of coordinates, which in each case that the notation $\tilde{f}$ is used, is either to be given explicitly or to be understood. We call $\tilde{f}$ the *coordinate expression for* f (in the system (p_α, U_α) and (p_β, V_β)) and we say briefly that $\tilde{f}$ expresses f. If $f : M \to R^{n'}$, then $\tilde{f} = f p_\alpha^{-1}$ expresses f. If $f : R^n \to M'$ then $p'_\beta f = \tilde{f}$ expresses f, and if $f : R^n \to R^{n'}$, then $\tilde{f} = f$ and f is already expressed in coordinates. (Cf. the remark immediately preceding Definition 3.) We note the useful fact that if $f : M \to M'$, $g : M' \to M''$ then, as the equation

$p_\gamma'' g f p_\alpha^{-1} = p_\gamma'' g p_\beta'^{-1} p_\beta' f p_\alpha^{-1}$ shows, the *composition* of two C^∞ mappings f and g is C^∞ and is expressible in coordinates as the composition of the coordinate expressions of f and g.

We can now define the notion of C^∞ homeomorphism of manifolds. It is as significant for C^∞ manifolds as the notion of homeomorphism is for topological spaces.

4. Definition. If $f: M \to M'$ is $1-1$ and onto, and if f and f^{-1} are C^∞, we say that f is a C^∞ homeomorphism and that M and M' are C^∞ homeomorphic.

Thus, C^∞ homeomorphic manifolds are topologically equivalent and furthermore have the same C^∞ properties. A standard way of obtaining a manifold C^∞ homeomorphic to a given manifold M is to add more coordinate systems to the given family of coordinate systems of M.

For example, if P_0 is any point, we may add a coordinate system (p, U) to M such that $p(P_0) = 0$ and $p(U)$ is a sphere centered at 0. To see this, let (p_1, U_1) be any coordinate system of M containing P_0. Let $p_1(P_0) = x_0$ and, noting that $p_1(U_1)$ is open, let S_ε be a sphere centered at x_0 with $S_\varepsilon \subseteq p_1(U_1)$. Now let t be the translation of R^n sending x_0 into 0. We add the coordinate system $U = p_1^{-1}(S_\varepsilon)$, $p = tp_1|U$. Then (p, U) has the required properties. Since t is surely C^∞, we obtain an equivalent manifold. In an entirely similar way, by using magnifications, we may construct a coordinate system (p, U) such that $p(U)$ is the unit sphere centered at 0 with $p(P_0) = 0$.

We have defined C^∞ manifolds and C^∞ mappings and functions, and now wish to differentiate them. The proper generalization to a manifold, as to n-dimensional Euclidean space, of the notion of differentiation is given by the directional derivative. Elaborating this idea, we will arrive at the concept of a tangent vector to M at a point P_0.

5. Definition. Let $P(t)$ be a (C^∞) mapping of the real variable t in some open interval with values in M. Then we say that $P(t)$ is a *curve* in M. If $P(t_0) = P_0$ we say that the curve $P(t)$ *passes through* P_0 at $t = t_0$. If $P(0) = P_0$ we often simply say that $P(t)$ *passes through* P_0.

6. Definition. Let f be a real-valued function on M and let $P_0 \in M$. Suppose

$$\left.\frac{df(P(t))}{dt}\right|_{t=0} = 0$$

for every curve $P(t)$ passing through P_0 ($P(0) = P_0$). Then we say that f *is horizontal at* P_0, or that $df|_{P=P_0} = 0$.

Definition 6 allows us a conveniently coordinate-free definition of the notion of a tangent vector.

7. Definition. A tangent vector λ at $P_0 \in M$ is a linear functional defined on the space of all real (C^∞) functions f on M and having the property that $\lambda(f) = 0$ if f is horizontal at P_0.

The tangent vectors at P_0 may be added and multiplied by scalars in an obvious way and thus form a vector space (of linear functionals).

8. Definition. The tangent space $\Lambda(P_0)$ of M is the vector space of all tangent vectors λ at P_0.

The following lemmas show the sense in which the notion of a tangent vector is a purely local notion.

9. Lemma. If $\lambda \in \Lambda(P_0)$ and if f and g are real valued functions on M which agree in some neighborhood of P_0, then $\lambda(f) = \lambda(g)$.

Proof. $f(P) - g(P) = 0$ near P_0. Therefore, by Definition 6. $f - g$ is certainly horizontal at P_0. Hence $\lambda(f - g) = 0$ and the result is proved. Q.E.D.

10. Lemma. If $P_0 \in M$ and f is defined, real-valued, and C^∞ in a neighborhood of P_0, then there is a real C^∞ function $\hat{f}$ defined on all of M which agrees with f in some neighborhood of P_0.

Proof. Let $\tilde{f}(x)$ be the coordinate expression for f in some coordinate system (p, U). With no loss in generality, we may assume that $p(U) = \{x \in R^n | \, |x| < 1\}, p(P_0) = 0$, so that $\tilde{f}(x)$ is defined for $|x| < 1$ and the coordinate of P_0 is 0. By a familiar result of ordinary analysis, there is a real C^∞ function $\tilde{t}(x)$ defined for $|x| < 1$ and satisfying

$$\tilde{t}(x) \equiv 1 \text{ for } |x| < 1/3, \qquad \tilde{t}(x) \equiv 0 \text{ for } |x| > 2/3.$$

Define $t(P)$ for $P \in U$ by $t(P) = \tilde{t}(p(P))$, and then define the required $\hat{f}$ by the formulas

$$\hat{f}(P) = f(P)\, t(P) \qquad P \in U$$

$$\hat{f}(P) = 0 \qquad P \notin U$$

It is readily verified that $\hat{f}(P)$ is C^∞ at all points of M and $\hat{f} = f$ in a neighborhood of P_0. Q.E.D.

Lemmas 9 and 10 show that a tangent vector λ at P_0 acts in a natural way on functions defined only in some neighborhood of P_0. This permits the complete analysis of the tangent space $\Lambda(P_0)$ with the help of a coordinate system about P_0.

Our first step toward this goal must be to interpret Definition 6 in coordinates. In order to facilitate computation, we shall write equations "dropping higher powers." This method will be used occasionally in what follows and we review it here. If $f(x)$ is a real (C^∞) function of the real variable x near $x = 0$, and if k is any integer, then there is a *polynomial* $f_k(x) = a_0 + a_1x + \cdots + a_kx^k$ which agrees with f and all of its derivatives up through order k at $x = 0$. We write

$$f(x) = f_k(x) + \cdots = a_0 + \cdots + a_kx^k + \cdots$$

"to order k." Then the assignment $f \to f_k(x)$ is linear and every reasonable computational operation involving the first k derivatives of f at $x = 0$ may be performed in a natural way. In particular, we may differentiate to obtain

$$f'(x) = a_1 + \cdots + ka_kx^{k-1} + \cdots (= f_{k-1}(x) + \cdots)$$

and we may substitute $x = b_1t + b_2t^2 + \cdots + b_kt^k + (\cdots)$ to obtain the correct expression for $f(x(t))$ to order k after further dropping powers higher than k. Similar rules apply to a vector-valued function of the real variable x, as may be seen by working with the components of the vector values of f. Thus, if f is vector-valued, we may write

$$f(x) = a_0 + xa_1 + \cdots + x^ka_k + \cdots$$

to order k, where a_i are constant vectors. We may also generalize this notation to functions of n real variables. If $x = (x^1, \ldots, x^n)$ is a vector variable and f is real-valued, then $f_k(x)$ is the polynomial in $x^1, \ldots, x^n$ which agrees with f and all of its mixed derivatives through order k at $x = 0 = (0, \ldots, 0)$. We then write

$$f(x) = a_0 + a_1(x) + a_2(x, x) + \cdots + a_k(x, \ldots, x) + \cdots$$

where $a_i(x_1, \ldots, x_i)$ is a real-valued symmetric i-linear functional. Finally, if f is a vector-valued function of the vector x near $x = 0$, this last formula remains valid except that a_i is a vector-valued symmetric i-linear function. Similar series of powers of $(x - x_0)$ will sometimes be used when we work with $f(x)$ and its derivatives at $x = x_0$.

If f is real and $x = (x^1, \ldots, x^n)$, and if $f(x_0 + x) = a_0 + a(f; x) + \cdots$

then by definition $a_0 = f(x_0) = f(x_0^1, \dots, x_0^n)$ while if $a(f; x) = \sum a_i x^i$, we have by definition

$$\left.\frac{\partial f(x_0 + x)}{\partial x^i}\right|_{x=0} = \left.\frac{\partial f(x)}{\partial x^i}\right|_{x=x_0} = \left.\frac{\partial a(f; x)}{\partial x^i}\right|_{x=x_0} = a_i$$

Thus we have, more explicitly,

$$f(x_0 + x) = f(x_0) + \sum_{i=1}^{n} \left.\frac{\partial f(x)}{\partial x^i}\right|_{x=x_0} x^i + \cdots$$

and therefore the coefficients of $a(f; x)$ are the components of the gradient of f at $x = x_0$. Similarly, if $y = f(x)$ is vector-valued with $y = (y^1, \dots, y^m)$, $f(x) = (f^1(x), \dots, f^m(x))$, we have

$$f^i(x_0 + x) = f^i(x_0) + \sum_{j=1}^{n} \left.\frac{\partial f^i(x)}{\partial x^j}\right|_{x=x_0} x^j$$

and the coefficients of $a(f, x)$ are the entries of the Jacobian matrix of f at $x = x_0$.

Using these notations, we now interpret Definition 6 in coordinate form.

11. Lemma. Let f be real-valued on M and let (p, U) be a coordinate system containing P_0 with $p(P_0) = x_0$. Let $\tilde{f}$ be the coordinate expression for f and

$$\tilde{f}(x_0 + x) = a_0 + a(\tilde{f}; x) + \cdots$$

to first order. Then f is horizontal at P_0 if and only if $a(\tilde{f}; x) = 0$, or equivalently $\left.\frac{\partial \tilde{f}(x)}{\partial x^i}\right|_{x=x_0} = 0$ for all i.

The proof is trivial and is left to the reader. Note the following consequence: if $a(\tilde{f}; x) = 0$ in one coordinate system, the same is true in all coordinate systems. This follows from Lemma 11, since the definition of the phrase "f is horizontal at P_0" was coordinate-free.

12. Lemma. Let M be an n-dimensional manifold, $P_0 \in M$. Then $\Lambda(P_0)$ is an n-dimensional vector space. Let (p, U) be a coordinate system about P_0 with $p(P_0) = x_0$. Then for each $\lambda \in \Lambda(P_0)$ there is a unique $\tilde{\lambda} \in R^n$ (called its coordinate) satisfying

$$\lambda(f) = a(\tilde{f}; \tilde{\lambda}). \tag{1}$$

Here, $a(f; x)$ is the linear part of f at x_0:

$$\tilde{f}(x_0 + x) = c + a(\tilde{f}; x) + \cdots \tag{2}$$

The mapping $\lambda \to \tilde{\lambda}$ is linear, $1 - 1$, and onto.

Proof. Let $\lambda \in \Lambda(P_0)$. By Lemmas 9 and 10, we may regard λ as acting on all functions defined in the neighborhood U of P_0. If $\tilde{f}$ is the coordinate expression for f, we define $\lambda_1(\tilde{f}) = \lambda(f)$. Then λ_1, as a linear function defined on all functions in $p(U)$, in particular acts linearly on the space of functions of the form $a(x)$ where a is a linear real-valued functions on R^n. But any linear functional on the space of linear function on R^n is given by a unique vector $\tilde{\lambda}$ in R^n, i.e. there exists a $\tilde{\lambda} \in R^n$ such that

$$\lambda_1(a(x)) = a(\tilde{\lambda})$$

if $a(x)$ is a linear function on R^n. Now let $\tilde{f}$ be any function defined in $p(U)$, and let

$$\tilde{f}(x) = c + a(\tilde{f}; x - x_0) + \cdots$$

(for x near x_0). By Lemma 11, $\tilde{f}(x) - a(\tilde{f}; x)$ is horizontal at x_0. Hence $\lambda(f) = \lambda_1(\tilde{f}(x)) = \lambda_1(a(\tilde{f}; x)) = a(\tilde{f}; \tilde{\lambda})$. But this is equation (1).

It is now straightforward to show that the correspondence $\lambda \to \tilde{\lambda}$ is linear, $1 - 1$ and onto; we leave this concluding part of the proof to the reader. Q.E.D.

Note that equation (1) of the preceding lemma may be written in more explicit coordinate form as

$$\lambda(f) = a(\tilde{f}; \tilde{\lambda}) = \sum \frac{\partial \tilde{f}}{\partial x^i} \tilde{\lambda}^i \bigg|_{x = x_0},$$

where $\tilde{\lambda} = (\tilde{\lambda}^1, \ldots, \tilde{\lambda}^n)$.

We next generalize the notion of the Jacobian matrix of a collection of functions of several variables, to mappings of manifolds into manifolds.

13. Lemma. Let $g: M \to M'$, $g(P_0) = P_0'$, and let M and M' have dimensions n, n' respectively, and tangent spaces Λ and Λ' at P_0 and P_0' respectively. Let f' be a real-valued function on M' which is horizontal at P_0'. Then $f'g$ (i.e. the composed function defined by $f'g(P) = f'(g(P))$) is horizontal at P_0.

Proof. Let $P(t)$ be any curve through P_0. Then $gP(t)$ is a curve through P'_0. Hence, by hypothesis, $\frac{d}{dt} f'(gP(t))\Big|_{t=0} = 0$. But this is simply $\frac{d}{dt}(f'g)(P(t))\Big|_{t=0} = 0$ and since $P(t)$ was arbitrary through P_0, this proves the result. Q.E.D.

14. Definition. With g, M, M', P_0, P'_0, Λ, Λ' as in Lemma 13, we define $g_*: \Lambda \to \Lambda'$ by the formula

$$(g_*\lambda)(f') \doteq \lambda(f'g)$$

where f' is any real function on M'.

In view of Lemma 13 and the obvious linearity of the functional $g_*\lambda$, we see that $g_*\lambda \in \Lambda'$. Thus, if $g: M \to M'$, g_* maps the tangent space at any point $P_0 \in M$ into the corresponding tangent space at $P'_0 = g(P_0)$. In stating the following lemma, we introduce another manifold M'', of dimension n'', and a mapping $h: M' \to M''$ with $h(P'_0) = P''_0$. We let Λ'' be the tangent space to M'' at P''_0.

15. Lemma. g_* is a linear function of λ. $(hg)_* = h_*g_*$. Furthermore, $(\text{Identity})_* = \text{Identity}$.

The proof uses only Definition 14 and is trivial.

The next lemma shows how g_* is to be computed in coordinates.

16. Lemma. Let $g: M \to M'$ and let (p, U), (q, V) be coordinate systems containing $P_0 \in M$ and $P'_0 = g(P_0) \in M'$. Let $\tilde{g}$ be the coordinate representation of g and let $\tilde{\lambda}$, $\tilde{\mu}$ be the coordinates of the tangent vectors λ, μ in Λ and Λ' respectively. Then if P_0, P'_0 have coordinates x_0, y_0 respectively and if

$$y = \tilde{g}(x_0 + x) = y_0 + a(\tilde{g}; x) + \cdots$$

then

$$\widetilde{g_*(\lambda)} = a(\tilde{g}, \tilde{\lambda})$$

Proof: If the function $f(P')$ is represented by $\tilde{f}(y)$, and

$$\tilde{f}(y_0 + y) = c + a(\tilde{f}; y) + \cdots$$

then fg is represented by $\widetilde{fg}$, and

$$\begin{aligned}\widetilde{fg}(x_0 + x) &= \tilde{f}(y_0 + a(\tilde{g}; x) + \cdots) \\ &= c + a(\tilde{f}; a(\tilde{g}; x)) + \cdots\end{aligned}$$

If we let $\mu = g_*(\lambda)$, then by Definition 14, we have $\mu(f) = \lambda(fg)$. Using Lemma 12 and the above, we may write this as

$$a(\tilde{f}; \tilde{\mu}) = a(\tilde{f}; a(\tilde{g}, \tilde{\lambda}))$$

Since f is arbitrary, this implies

$$\widetilde{g^*(\lambda)} = \tilde{\mu} = a(\tilde{g}, \tilde{\lambda})$$

which is the required result. Q.E.D.

In more explicit coordinate form the preceding lemma tells us that the matrix of the transformation g_* which sends λ into μ is the Jacobian matrix of $y = \tilde{g}(x)$ at $x = x_0$:

$$\tilde{\mu}^i = \sum_{j=1}^{n} \left.\frac{\partial \tilde{g}^i(x)}{\partial x^j}\right|_{x=x_0} \tilde{\lambda}^j, \quad i = 1, \ldots, n'$$

where $\tilde{\lambda} = (\tilde{\lambda}^1, \ldots, \tilde{\lambda}^n)$ and $\tilde{\mu} = (\tilde{\mu}^1, \ldots, \tilde{\mu}^{n'})$. This is so because the matrix of the linear transformation $x \to a(\tilde{g}; x)$ is the Jacobian matrix of the transformation $y^i = \tilde{g}^i(x^1, \ldots, x^n)$ $(i = 1, \ldots, n')$ at $x = x_0$.

Using Lemma 16, we may easily prove the following coordinate-free version of the ordinary implicit function theorem.

17. Lemma. Let $g: M \to M'$ be a C^∞ homeomorphism. Then g_* is $1-1$ and onto for each $P_0 \in M$. (In particular, $n = n'$.) Conversely, let $g: M \to M'$ be $1-1$ and onto and let g_* be $1-1$ and onto at each $P_0 \in M$. Then g^{-1} is C^∞ and thus g is a C^∞ homeomorphism.

Proof. If g is a C^∞ homeomorphism, we have gg^{-1} = identity and $g^{-1}g$ = identity. Hence $g_*(g^{-1})_*$ = identity and $(g^{-1})_* g_*$ = identity. But these equations show that g_* is 1–1 onto with $(g_*)^{-1} = (g^{-1})_*$ at each point $P_0 \in M$.

Conversely, suppose that g is $1-1$ onto and that g_* is $1-1$ onto at each $P_0 \in M$. By Lemma 16, this shows that the Jacobian matrix of $\tilde{g}$ in any coordinate system is non-singular. Hence by the implicit function theorem, $\tilde{g}^{-1}$ exists and is C^∞ in some neighborhood of x_0'. But clearly $\widetilde{g^{-1}} = \tilde{g}^{-1}$. Thus g^{-1} is C^∞ in the neighborhood of each point $P_0' \in M'$. Hence g is a C^∞ homeomorphism. Q.E.D.

In later sections, we shall also have use for the following local generalizations of Lemma 17.

18. Lemma. Let $g:M \to M'$ be a C^∞ mapping. Let $P_0 \in M$, and $P'_0 = g(P_0) \in M'$.

(a) Suppose that g_* maps the tangent space $\Lambda(P_0)$ onto the tangent space $\Lambda(P'_0)$. Then, if $U \subseteq M$ is any neighborhood of P_0, $g(U)$ contains a neighborhood of P'_0.

(b) Suppose that g_* maps the tangent space $\Lambda(P_0)$ $1 - 1$ into the tangent space $\Lambda(P'_0)$. Then there is a neighborhood U of P_0 in M such that g is $1 - 1$ in U.

(c) Suppose that g_* maps the tangent space $\Lambda(P_0)$ $1 - 1$ onto the tangent space $\Lambda(P'_0)$. Then there are neighborhoods $U \subseteq M$ and $U' \subseteq M'$ of P_0 and P'_0 respectively such that g is a C^∞ homeomorphism from U onto U'.

Proof. (a) Let M and M' be n and n' dimensional respectively. Introduce coordinates in such a way that P_0 has x_0 as coordinate and P'_0 has x'_0 as coordinate. Write

$$\tilde{g}(x_0 + x) = x'_0 + a(\tilde{g}; x) + \cdots$$

Then by Lemma 16 and by hypothesis, the Jacobian transformation $x \to a(g; x)$ maps R^n onto $R^{n'}$. Hence there exists an n'-dimensional subspace $R_1 \subset R^n$, such that $x \to a(\tilde{g}; x)$ is a $1 - 1$ linear transformation of R_1 onto R^{n}. It follows by the ordinary implicit function theorem that if ε is sufficiently small the restriction of g to the set $\{x_0 + x \mid |x| < \varepsilon, x \in R_1\}$ is a homeomorphism whose range is a neighborhood of x'_0. Passing back to the manifolds, we have our result.

The proofs of (b) and (c) are similarly dependent on the implicit function theorem, and are left to the reader. Q.E.D.

We now wish to define the sense in which a curve through P_0 determines a tangent vector at P_0. Suppose $P = P(t)$, with $P(t_0) = P_0$. Let f be a real C^∞ function on M. Then, by Definitions 6 and 7 the linear functional which sends f into

$$\frac{d}{dt} f(P(t))\bigg|_{t=t_0}$$

is a tangent vector at $P_0 = P(t_0)$. This justifies the following definition and notation.

19. Definition. Let $P(t)$ be a C^∞ curve with $P(t_0) = P_0$. Then $\dfrac{dP}{dt}\bigg|_{t=t_0}$ is the tangent vector λ at P_0 defined by

$$\lambda(f) = \frac{d}{dt} f(P(t))\bigg|_{t=t_0}$$

If $\lambda = \left.\frac{dP}{dt}\right|_{t=t_0}$, we say that the curve $P(t)$ represents λ at $t = t_0$.

The following lemma interprets Definition 19 in coordinates, and gives an extra bit of information besides.

20. Lemma. (a) If $P(t)$ is a curve in M, we have

$$\left.\widetilde{\frac{dP}{dt}}\right|_{t=t_0} = \left.\frac{d}{dt}\tilde{P}(t)\right|_{t=t_0}$$

(b) If λ is a vector tangent to M at P_0, there exists a curve $P(t)$ with $P(t_0) = P_0$ and $\lambda = \left.\frac{dP}{dt}\right|_{t=t_0}$.

Proof. Let x_0 be the coordinate of P_0 and let $\tilde{f}$ represent the real function f on M:

$$\tilde{f}(x_0 + x) = c + a(\tilde{f}; x) + \cdots$$

Now let $P(t)$ be a curve passing through P_0 at $t = t_0$, and suppose $\left.\frac{d}{dt}\tilde{P}(t)\right|_{t=t_0} = \tilde{\mu}$ and that $\tilde{\mu}$ is the coordinate of the tangent vector μ. Then

$$\tilde{P}(t) = x_0 + (t - t_0)\,\tilde{\mu} + \cdots \qquad (t \text{ near } t_0)$$

$$\left.\frac{dP}{dt}\right|_{t=t_0}(f) = \left.\frac{d}{dt} f(P(t))\right|_{t=t_0} = \left.\frac{d}{dt}\tilde{f}(\tilde{P}(t))\right|_{t=t_0}$$

$$= \left.\frac{d}{dt}\left\{c + a(\tilde{f}; (t - t_0)\tilde{\mu}) + \cdots\right\}\right|_{t=t_0}$$

$$= a(\tilde{f}; \tilde{\mu}) = \mu(f)$$

Thus $\mu = \left.\frac{dP}{dt}\right|_{t=t_0}$ and $\tilde{\mu} = \left.\frac{d\tilde{P}}{dt}\right|_{t=t_0}$. This completes the proof since clearly any $\tilde{\mu}$ can occur. Q.E.D.

The following easy lemma gives a useful consistency property of our notations.

21. Lemma.

$$g_*\left(\left.\frac{dP}{dt}\right|_{t=0}\right) = \left.\frac{d}{dt}(gP)\right|_{t=0}$$

Proof.
$$g_*\left(\frac{dP}{dt}\bigg|_{t=0}\right)(f') = \left(\frac{dP}{dt}\bigg|_{t=0}\right)(f'g)$$
$$= \frac{d}{dt} f'g(P(t))\bigg|_{t=0}$$
$$= \frac{d}{dt}(gP)\bigg|_{t=0}(f') \quad \text{Q.E.D.}$$

Remark. In brief, $\frac{d}{dt}(gP) = g_*\left(\frac{dP}{dt}\right)$.

The preceding notation has been chosen so that it generalizes notation used in the special case that M is a vector space or $M = R^n$. (Cf. the remark following Definition 2.) If M is a vector space, a tangent vector at $P_0 \in M$ is given by a vector $\lambda \in M$ and if f is any real-valued function of M we have $\lambda(f) = \frac{d}{dt} f(P_0 + \lambda t)\Big|_{t=0}$. This convention has the effect of identifying λ and its coordinate $\tilde{\lambda}$ in case M is a vector space. If $P = P(t)$ is a curve in M, Lemma 20 shows that the tangent vector dP/dt in the sense of Definition 19 is simply the ordinary derivative of the vector-valued function $P(t)$ of the real variable t.

We shall have occasion to consider differential equations of the form $\frac{dP}{dt} = \lambda(P, t)$. We must therefore define the notion of a C^∞ tangent vector field.

22. Definition. Let $\lambda(P)$ be a tangent vector at P for each $P \in M$. $\lambda(P)$ is called a C^∞ vector field if $\lambda(P)(f)$ is a C^∞ function of P for every real (C^∞) f on M.

23. Lemma. $\lambda(P)$ is C^∞ if and only if in some coordinate system $\tilde{\lambda}(x)$ is a C^∞ function of x (here, x is the coordinate of P).

Proof. If f is a real function on M, and if $\tilde{f}(y^1, \ldots, x^n)$ represents f and $\tilde{\lambda}(x^1, \ldots, x^n) = (\tilde{\lambda}^1(x^1, \ldots, x^n), \ldots, \tilde{\lambda}^n(x^1, \ldots, x^n))$ represents λ, then by the remark immediately following Lemma 12 we have

$$\lambda(P)(f) = \sum_{i=1}^{n} \tilde{\lambda}^i(x) \frac{\partial \tilde{f}}{\partial x^i}.$$

Hence if $\tilde{\lambda}$ is C^∞ in any coordinate system, $\lambda(P)$ is C^∞.

The converse is equally trivial and is left to the reader.

We may, in a similar manner, define the notion of a C^∞ vector field in some open set of M, and also a tangent vector field $\tilde{\lambda}(P, t^1, \ldots, t^k)$ (parametrized by $t^1, \ldots, t^k$) which is simultaneously C^∞ in $(P, t^1, \ldots, t^k)$. In this latter case, $\tilde{\lambda}(x^1, \ldots, x^n, t^1, \ldots, t^k)$ is C^∞ as a function of $x^1, \ldots x^n, t^1, \ldots, t^k$. We leave the details to the reader. The following lemma is presumably anticipated by the reader.

24. Lemma. Let $\lambda(P, t)$ be a tangent vector field simultaneously C^∞ in P and t. Then the equation $\dfrac{dP}{dt} = \lambda(P, t)$, $P(t_0) = P_0$ has a unique C^∞ solution $P = P(t)$ near $t = t_0$.

Proof. In coordinates, this amounts to solving $\dfrac{dx}{dt} = \tilde{\lambda}(x, t), x(t_0) = x_0$ where $\tilde{\lambda}$ is C^∞. This is known to have a C^∞ solution $x(t)$ near $t = t_0$. Q.E.D.

25. Lemma. If $\lambda(P, t, a^1, \ldots, a^k)$ is simultaneously C^∞ in $(P, t, a_1, \ldots, a_k)$, then the solution of

$$\frac{dP}{dt} = \lambda(P, t, a^1, \ldots, a^k); \; P(t_0) = P_0,$$

which may be denoted $P = P(t, a^1, \ldots, a^k)$ to denote dependence on $(a^1, \ldots, a^k)$, is defined for t near t_0 and for all $a = (a^1, \ldots, a^k)$ in a sufficiently small open set, and is C^∞ as function of $(t, a^1, \ldots, a^k)$.

Proof. The proof is clear and is based on the analogous result in differential equations. Q.E.D.

We now consider two of the usual ways of obtaining new structures from old: the product structure and the sub-structure. The first is automatic, the second is surprisingly intricate.

26. Definition. Let M_1, M_2 be manifolds of dimensions n_1, n_2 respectively and with coordinate covers (p^1_α, U_α) and (p^2_β, V_β) respectively. Then $M_1 \times M_2$, *the product manifold*, is defined as the manifold whose topology is the usual product topology of M_1 and M_2 and whose coordinate covering is $(p_{\alpha\beta}, U_{\alpha\beta})$ where $U_{\alpha\beta} = U_\alpha \times V_\beta$ and $p_{\alpha\beta} \colon U_{\alpha\beta} \to R^{n_1+n_2}$ is defined by the formula $p_{\alpha\beta}(P, Q) = (p^1_\alpha P, p^2_\beta Q)$ where $P \in U_\alpha$, $Q \in V_\beta$.

There are many obvious results concerning the connection between the product space and its factors. For example, the tangent space at (P_1, P_2) of $M_1 \times M_2$ is, in a natural way, the direct sum of the tangent spaces

$\Lambda_1(P_1)$ of M_1 and $\Lambda_2(P_2)$ of M_2. Similarly if $f: M \to M_1' \times M_2'$ has components (f_1, f_2), then f_* has components (f_{1*}, f_{2*}). We shall not go into these matters in any more detail.

One would expect the definition of submanifold to begin, "A submanifold M_1 of M is a subspace M_1 of M with coordinate mappings..." This is not the definition adopted. In fact, we shall have submanifolds whose subspace topology is not even locally Euclidean. Diagram 1 below gives an example of the submanifold of the plane, which if the subspace topology were adopted would not even be locally Euclidean. A neighborhood of P_0 in the topology which we actually intend to use is shaded in the diagram. The sequence P_n does not approach P_0 in the topology of this submanifold.

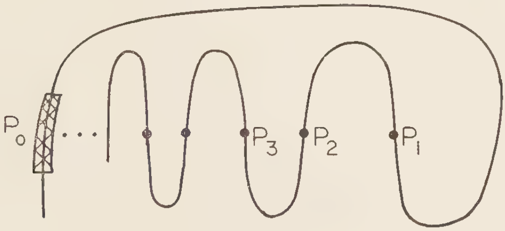


Diagram 1: A Submanifold of the Plane

An important example of a pathological submanifold, which is more closely related to the subject matter of these notes, is as follows. Let T^2 be the 2-dimensional torus consisting of all couples (x, y), in which (x, y) and (x', y') are identified when $x - x'$ and $y - y'$ are integers. Thus, $T^2 = R^2/N$, when R^2 is taken as the 2-dimensional additive group, and where the discrete subgroup $N = \{(n, m) | \; n \text{ and } m \text{ integers}\}$ is a normal subgroup. T^2 is clearly a Lie group. (See Theorem 2.11 for the formal definition of such a quotient Lie group.) Now, if θ is an irrational number, the mapping $f(t) = (t, \theta t)$ is a $1 - 1$ homomorphic mapping of the reals into the torus. The image of the reals is *one-dimensional, yet everywhere dense* in the torus. (See diagram 2 below.)

For an n-torus T^n, the mapping $f(t) = (\theta_1 t, \ldots, \theta_n t)$ similarly gives rise to a one-dimensional subgroup which is dense in T^n provided $\theta_1, \theta_2, \ldots, \theta_n$ are linearly independent over the rationals.

The possible occurrence of pathology of this kind lies behind the fussiness with which the forthcoming definitions and results concerning submanifolds and (in sections 2 and 6) subgroups, must be formulated.

27. Definition. Let $g: M \to M'$. g is said to be *regular at* P_0 if g_* is $1 - 1$ at P_0. Similarly, g is *regular* if it is regular at each of its points. If g is not regular at P_0 it is called *singular* there.

We now define a submanifold M_1 of M, with the help of an external manifold N, which is to be thought of as a replica of M_1.

28. Definition. Let M and N be manifolds, let $M_1 \subseteq M$ and let $g: N \to M$ be regular and $1 - 1$, and let g map N onto M_1. If M_1 is made into a manifold by transferring the topology and C^∞ structure from N via g, then M_1 is called a submanifold of M (via N, g). Explicitly, the open sets of M_1 are the images $g(U)$ where U is open in N and the coordinate cover of N consists of $(p_\alpha g^{-1}, gU_\alpha)$ where (p_α, U_α) constitute the coordinate cover of N.

We have set things up so that N and M_1 are obviously (C^∞) homeomorphic. For if we define $g_1: N \to M_1$ by the formula $g_1(P) = g(P)$ then g_1 is clearly a homeomorphism. The inclusion map $i: M_1 \to M$, defined by $i(P) = P$ is C^∞, since $i = gg_1^{-1}$. From this, it follows that the topology of M_1 is stronger than (or equal to) the subspace topology of M_1, since i is continuous. It also follows that if $f: M \to M'$ is C^∞, then $f|M_1$ is also C^∞.

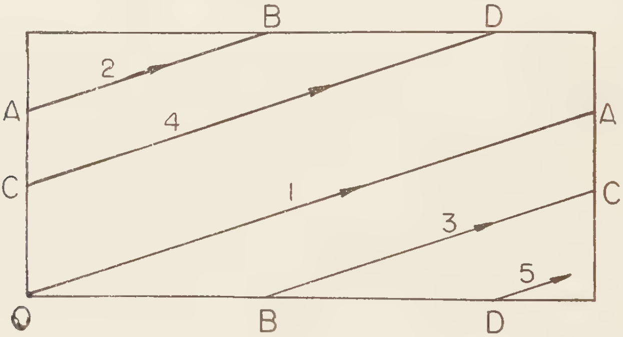


Diagram 2.

Indeed $f|M_1 = fi$. We shall now show that conversely any C^∞ mapping f defined on a submanifold is, locally on M_1, the restriction of a C^∞ mapping $\hat{f}$ on M.

29. Lemma. Let M_1 be a submanifold of M via (N, g). Let $f: M_1 \to M'$, and let $P_0 \in M_1$. Then there exists a mapping $\hat{f}$: $M \to M'$ and a neighborhood of U of P_0 in M_1 such that $\hat{f} = f$ in U. Briefly, any C^∞ mapping on M_1 is locally the restriction of a C^∞ mapping on M.

Proof: We work with the replica N of M_1. Thus, let $f':N \to M'$, and let $g:N \to M$ be $1-1$ and regular, while $P_0' \in N$. We wish to construct an $\hat{f}:M \to M'$ such that $\hat{f}g = f'$ near P_0'. Let $x = (x^1, \dots, x^r)$ be coordinates for N, $y = (y^1, \dots, y^n)$ be coordinates for M, and $z = (z^1, \dots, z^{n'})$ be coordinates for M'. Let the coordinate expression of the maps g and f' be $y = \tilde{g}(x)$ and $z = \tilde{f}'(x)$ respectively, so that

$$y^i = \tilde{g}^i(x^1, \dots, x^r) \qquad i = 1, \dots, n \tag{1}$$

$$z^j = \tilde{f}'^j(x^1, \dots, x^r) \qquad j = 1, \dots, n'. \tag{2}$$

Because g is regular, the Jacobian matrix $\dfrac{\partial \tilde{g}^i}{\partial x^k}$ has rank r near $x = x_0$. By the implicit function theorem we may solve a subset of r of the equations (1) for the x's in terms of the corresponding y's—say, $y^1, \dots, y^r$:

$$x^k = h^k(y^1, \dots, y^r) \qquad k = 1, \dots, r. \tag{3}$$

We now define $\hat{f}:M \to M'$ locally by the equations:

$$z^j = z^j(y) = \tilde{f}'^j(h^1(y^1, \dots, y^r), \dots, h^r(y^1, \dots, y^r)). \tag{4}$$

Then clearly $\hat{f}g = f'$ near P_0'. This proves the assertion of the present lemma, except that we wish to have $\hat{f}$ defined on all of M. But this is clear from Lemma 10. Q.E.D.

If M_1 is a submanifold of M, the tangent space $\Lambda_1(P_0)$ of M_1 at $P_0 \in M_1$ is included in a natural way in the tangent space $\Lambda(P_0)$ of M at P_0.

30. Lemma. Let M_1 be a submanifold of M, let $P_0 \in M_1$, and let $\Lambda_1 = \Lambda_1(P_0)$, $\Lambda = \Lambda(P_0)$ be the tangent spaces of M_1 and M respectively at P_0. Let i be the inclusion map $i:M_1 \to M$. Then $i_*:\Lambda_1 \to \Lambda$ is $1-1$. Hence Λ_1 is embedded naturally, by means of i_*, in Λ.

Proof. We already know, from the remarks following Definition 28, that the map i is C^∞. By Definition 14, it follows that if $\lambda_1 \in \Lambda_1$, and if f is any real-valued function on M, we have

$$(i_*\lambda_1)f = \lambda_1(fi) = \lambda_1(f|M_1)$$

since $fi = f|M_1$. We must show that i_* is $1-1$ from Λ_1 into Λ. Suppose $\lambda_1 \neq 0$. We may find a real-valued function f on M_1 such that $\lambda_1(f) \neq 0$. By the preceding lemma, there is a function $\hat{f}$ on M such that $\hat{f} = f$ in some neighborhood U_1 of P_0 in M_1. Hence $f = \hat{f}|U_1$. By Lemma 9, and the above equations, we have

$$(i_*\lambda_1)\hat{f} = \lambda_1(\hat{f}|M_1) = \lambda_1(f) \neq 0.$$

Hence $i_*\lambda_1 \neq 0$ and i is $1 - 1$. Q.E.D.

Remark. We embed Λ_1 in Λ by letting tangent vectors to M_1 act on functions on M by first restricting these functions to M_1. We note that if $P(t)$ is a (C^∞) curve in M_1, it is naturally identified with a curve in M (i.e., $iP(t)$ is a curve in M), and the respective tangent vectors are also identified. This follows formally from Lemma 21: $\frac{d}{dt}(iP(t)) = i_*\left(\frac{dP}{dt}\right)$.

Warning: Suppose $g: M' \to M$ is C^∞, and M_1 is a submanifold of M such that $g(M') \subseteq M_1$. It is not necessarily true that "g": $M' \to M_1$ is C^∞. As an extreme example, M_1 may have the discrete topology. On the other hand, if $g: M' \to M_1$ is C^∞, the "g": $M' \to M$ is also C^∞, for this latter function is merely ig.

The following lemma gives a useful, commonly used example of a submanifold.

31. Lemma. Let M_1 be an open set in M. Then M_1 may be made into a submanifold in the following standard way. Letting (p_α, U_α) be a coordinate cover for M, take $(p_\alpha|U_\alpha \cap M_1, U_\alpha \cap M_1)$ as a coordinate cover for M_1, omitting those α for which $U_\alpha \cap M_1$ is empty. In this way M_1 becomes a manifold of the same dimension as M, and moreover has the subspace topology.

Proof. It is trivial that this definition makes a M_1 into a manifold of the same dimension as M and that M_1 has the subspace topology. In terms of Definition 28, we have merely to take $N = M_1$ and $g =$ the inclusion map of M_1 into M, and this shows that M_1 is a submanifold. Q.E.D.

Remark. If M_1 is an open subset of M, M_1 will always be taken as a submanifold of M in the sense of Lemma 31, unless otherwise specified. An important example of submanifolds of this special sort arises as follows. If M_1 is the connected component of a point $P_0 \in M$, is clearly M_1 open, since M is locally Euclidean. We therefore see that the connected components of a manifold M are submanifolds of the same dimension. The reader may also convince himself that in this situation the warning immediately preceding Lemma 31 does not apply.

2. Lie Groups: Local and Global Properties

A Lie group is a manifold which is also a group. In studying Lie groups we may therefore combine topology, analysis, and algebra. The present section is chiefly topological, and aims at showing the sense in which the global study of a Lie group may be reduced to its local study. In the next sections we shall go even further, and show that the study of the local structure can be reduced to the study of the infinitesimal structure. But rather than anticipate what is to come later, let us begin with our formal development.

1. Definition. A Lie group G is a C^∞ manifold endowed with a group structure in which multiplication and the taking of inverses are C^∞ operations. More precisely, we require that the mapping $(g, h) \to gh$ is C^∞ from $G \times G$ into G, and that the mapping $g \to g^{-1}$ is C^∞ from G into G.

It is an immediate consequence of the definition, that the inverse function $h = g^{-1}$ is a C^∞ homeomorphism of G onto G. For its inverse function $g = h^{-1}$ is C^∞. For fixed g_0, g_1, the functions $h_1(g) = g_0 g$, $h_2(g) = g g_0$ and $h_3(g) = g_0 g g_1$ are also C^∞ homeomorphisms of G onto G because their inverse functions are clearly of the same type, hence C^∞.

In what follows, we shall use the letter e to denote the group identity. Group elements will be denoted g, h, and k, possibly with subscripts. If A and B are sets in G, AB will denote the set of all products ab $(a \in A, b \in B)$, while gB will denote all products gb $(b \in B)$ etc. The symbol A^{-1} will denote the set all elements $a^{-1}(a \in A)$.

2. Lemma. Let G be a connected Lie group, U a neighborhood of the identity e. Then G is generated by U; i.e. any $g \in G$ is a (finite) product $g = g_1 \cdots g_n$ with $g_i \in U$, with n depending on U and g.

Proof. Since U^{-1} is the image of U under the homeomorphism $g \to g^{-1}$, U^{-1} is open and also contains $e = e^{-1}$. We shall work with the smaller neighborhood $V = U \cap U^{-1}$, which has the convenient property $V = V^{-1}$. Let $[V]$ be the set of products of the form $g_1 \cdots g_n$ with $g_i \in V$. We shall prove that $[V] = G$ by showing that $[V]$ is open and closed, and invoking the connectedness of G.

To show that $[V]$ is open, we let $g_0 = g_1 \cdots g_n \in [V]$. Then $g_0 V$ is open since left multiplication by g_0 is a homeomorphism. It is clearly contained in $[V]$ since it contains only elements of the form $g_1 \cdots g_n g_{n+1}$ with $g_{n+1} \in V$. Plainly, $g_0 = g_0 e \in g_0 V$. Thus $g_0 \in g_0 V \subseteq [V]$, which proves $[V]$ is open.

To show that $[V]$ is closed, let $h \in \overline{[V]}$. Then hV is an open set containing h, and it therefore meets $[V]$: $hV \cap [V] \neq \phi$. Let $g_0 \in hV \cap [V]$ so that $hh_1 = g_0 = g_1 \cdots g_n$ where $h_1, g_1, \ldots, g_n \in V$. Then

$$h = g_1 \cdots g_n h_1^{-1} \in [V],$$

and $[V]$ is closed. Q.E.D.

3. Lemma. Let G be a Lie group and let G_0 be the (connected) component of G containing the identity e. Then G_0 is an open invariant subgroup of G, and a Lie group, and the components of G are the cosets of G_0.

Proof. Since G is locally connected, it follows that G_0 is open. We have seen in lemma 1.31 how G_0 inherits the C^∞ structure of G. Thus, G_0 is a C^∞ manifold. Hence in order to show that it is a Lie group it suffices merely to show that it is closed under multiplication and the taking of inverses. For then the restriction of these operations to G_0 will be C^∞. Now, G_0^{-1}, as homeomorph of G_0 in G is connected. Hence, since $e = e^{-1} \in G^{-1} \cap G$, it follows that $G_0 \cup G_0^{-1}$ is connected. Hence $G_0 \cup G_0^{-1} = G_0$ and $G_0^{-1} \subseteq G_0$ so that G_0 is closed under the taking of inverses. Next, note that if G_0G_0 is the set of all products gh ($g \in G_0$, $h \in G_0$), then G_0G_0 is connected since it is the image of the connected set $G_0 \times G_0$ under the multiplication mapping. Since $e = ee \in G_0G_0$, we have $G_0G_0 \subseteq G_0$ and therefore G_0 is closed under multiplication. Thus, G_0 is a Lie group.

Since for any x, $e = xex^{-1} \in xG_0x^{-1}$ and xG_0x^{-1} is connected, we have $xG_0x^{-1} \subseteq G_0$, so that G_0 is an invariant subgroup.

Finally the left cosets xG_0 are all connected, since left-multiplication is a homeomorphism. They are maximal connected subsets of G since if $xG_0 \subset H$ properly and H is connected, then $G_0 \subset x^{-1}H$ properly and $x^{-1}H$ is connected, and this contradicts the definition of G_0. Q.E.D.

4. Definition. Let G, G' be Lie groups, and let $f: G \to G'$. f is called a *Lie group homomorphism*, or simply a *homomorphism* if f is a group homomorphism [$f(gh) = f(g)f(h)$] and f is C^∞. If further f is a homeomorphism, then it is called an isomorphism. If f is a mapping defined on some neighborhood U of e in G, if f maps U into G', and if $f(gh) = f(g)f(h)$ for g, h in some neighborhood V of e satisfying $VV \subseteq U$, then f is called a *local homomorphism*. If in addition, f is a homeomorphism of some neighborhood of e onto a neighborhood of the identity e' in G', then f is called a *local isomorphism*.

Clearly any homomorphism (isomorphism) of G into (onto) G' is a local homomorphism (isomorphism). We now show how these notions are related to each other.

5. Lemma. Let f be a C^∞ local homomorphism of the connected Lie group G into G'. Then there is at most one homomorphism $\hat{f}: G \to G'$ which agrees with f in a neighborhood of e.

Proof. The proof follows from Definition 4 and Lemma 2 since a homomorphism of a group is determined by the images of a set of generators for the group. Q.E.D.

6. Lemma. If $f: G \to G'$ is a homomorphism (not necessarily even continuous) and if f is C^∞ at e, then f is everywhere C^∞.

Proof. Let U be a neighborhood of e in which f is C^∞, and let $g_0 \in G$. Then $g_0 U$ is a neighborhood of g_0. We prove f is C^∞ in $g_0 U$. If $g \in g_0 U$, we have $f(g) = f(g_0 g_0^{-1} g) = f(g_0) f(g_0^{-1} g)$. This is C^∞ in $g_0 U$ since all mappings involved in this formula are C^∞. Q.E.D.

Lemma 5 shows the uniqueness of an extension of a local C^∞ homomorphism to a homomorphism in the large, while Lemma 6 shows that if the extension exists in the group theory sense it is automatically C^∞. The following theorem gives a condition for a local homomorphism to be extendable to a homomorphism in the large.

7. Theorem. Let f be a C^∞ local homomorphism of G into G', and let G be simply connected. Then there is a unique C^∞ homomorphism $\hat{f}: G \to G'$ such that $\hat{f} = f$ in a neighborhood of e.

Proof. Uniqueness follows from Lemma 5, while Lemma 6 shows that the homomorphism $\hat{f}$, if it exists, must be C^∞.

To show that $\hat{f}$ exists, we argue as follows. Let V_0 be a connected neighborhood of e such that $f(g_1 g_2) = f(g_1) f(g_2)$ for $g_i \in V_0$, and let $V \subseteq V_0$ be a connected neighborhood of e such that $V^{-1} V \subseteq V_0$. Let $g(t)$ $(a \leqslant t \leqslant b)$ be any continuous curve in G. We shall define an element $h = H[g(t)] \in G'$ depending on the curve $g(t)$ and having certain properties. (Ultimately, $H[g(t)]$ will be seen to be $\hat{f}(g(a))^{-1} \hat{f}(g(b))$.)

Call a partition $a_0 = a < a_1 < \cdots < a_n = b$ of $[a, b]$ *fine* (for the curve $g(t)$) if, for any subinterval $I_\nu = [a_{\nu-1}, a_\nu]$, the statement $t_1, t_2 \in I_\nu$ implies $g(t_1)^{-1} g(t_2) \in V_0$. It follows from the compactness of $[a, b]$ and the continuity of the group operations that there is a $\delta > 0$ such that if $|t_2 - t_1| < \delta$ then $g(t_1)^{-1} g(t_2) \in V_0$. Thus, fine partitions (for the curve

$g(t)$) exist. Put $\Delta g_\nu = g^{-1}(a_{\nu-1})\, g(a_\nu)$, so that $\Delta g_\nu \in V_0$. Now define $H[g(t)] = f(\Delta g_1) \cdots f(\Delta g_n)$.

We first show that H does not depend on the (fine) partition of $[a, b]$. Suppose that a point is added to the partition, i.e. that $[a_{\nu-1}, a_\nu]$ is replaced by $[a_{\nu-1}, c]$ and $[c, a_\nu]$. Then, since $\Delta g_\nu = g^{-1}(a_{\nu-1})\, g(a_\nu) = [g^{-1}(a_{\nu-1})\, g(c)]\, [g^{-1}(c)\, g(a_\nu)] = \Delta' g_\nu \cdot \Delta'' g_\nu$, and since f is a local homomorphism, we have $f(\Delta g_\nu) = f(\Delta' g_\nu)\, f(\Delta'' g_\nu)$. Thus H is unchanged if one point is added to the partition. Consequently, H does not depend on the partition at all, since any two fine partitions have a common refinement.

Next we note some properties of $H = H[g(t)]$.

$$H(g_0 g(t)) = H(g(t)). \tag{1}$$

$$H(g_1(t) * g_2(t)) = H(g_1(t))\, H(g_2(t)) \tag{2}$$

where $g_1(t) * g_2(t)$ is the product curve obtained by first traversing along g_1 and then along g_2, it being presumed here that the endpoint of g_1 coincides with the beginning point of g_2.

$$H(g(t)) = f(g(a))^{-1}\, f(g(b)) \tag{3}$$

if $g(t)$ is entirely within V. Finally

$$H(\text{constant}) = e' \tag{4}$$

where e' is the identity in G'. The proof of (1) is trivial, since $\Delta(g_0 g) = \Delta g$ because $(g_0 g_1)^{-1}\, (g_0 g_2) = g_1^{-1} g_2$. Statement (2) is also immediate since we may partition so that the common endpoint of g_1 and starting point of g_2 is in the partition. Statement (3) is equally clear since if $g(t)$ is entirely in V, $[a, b]$ is already a fine partition. Statement (4) is trivial.

We next show that H depends only on the homotopy class $[g(t)]$ of $g(t)$ (where the homotopy keeps end points fixed). For let $g_s(t)$ be a continuous mapping on the square $0 \leqslant s, t \leqslant 1$ into G which is a homotopy between $g_0(t)$ and $g_1(t)$. We have $g_s(0) = \text{constant}$, $g_s(1) = \text{constant}$. Let $0 \leqslant s_0 \leqslant 1$ and let $0 = a_0 < a_1 < \cdots < a_n = 1$ be a fine partition for $g_{s_0}(t)$. Then for s sufficiently close to s_0, these points also constitute a fine partition for $g_s(t)$; moreover $g_s(a_i)\, (g_{s_0}(a_i))^{-1}$ will lie in V_0. We now observe that it follows inductively for $1 \leqslant k \leqslant n$ that

$$\begin{aligned}
[*]\, & f(g_{s_0}(a_0)^{-1}\, g_{s_0}(a_1))\, f(g_{s_0}(a_1)^{-1}\, g_{s_0}(a_2)) \cdots f(g_{s_0}(a_{k-1})^{-1}\, g_{s_0}(a_k)) \\
&= f(g_s(a_0)^{-1}\, g_s(a_1))\, f(g_s(a_1)^{-1}\, g_s(a_2)) \cdots f(g_s(a_{k-1})^{-1}\, g_s(a_k)) \\
&\quad \times f(g_s(a_k)^{-1}\, g_{s_0}(a_k)).
\end{aligned}$$

Indeed, if $k = 1$, this follows from the formula $f(gh) = f(g)f(h)$ for $g, h \in V_0$ and the equation $g_s(0) = g_s(a_0) = g_{s_0}(a_0)$. We go from k to $k + 1$ by multiplying both sides of [*] by

$$f(g_{s_0}(a_k)^{-1}\, g_{s_0}(a_{k+1}))$$
$$= f(g_{s_0}(a_k)^{-1}\, g_s(a_k))\, f(g_s(a_k)^{-1}\, g_s(a_{k+1}))\, f(g_s(a_{k+1})^{-1}\, g_{s_0}(a_{k+1}))$$

Thus [*] holds for all k, and in particular $k = n$; therefore $H[g_s(t)] = H[g_{s_0}(t)]$ for all s near s_0, since $g_s(1) = g_s(a_n) = g_{s_0}(a_n)$. Since this holds for each s_0, we have of course $H[g_0(t)] = H[g_1(t)]$.

Now, G is simply connected. Hence, if g_1 and g_2 are any points of G, all curves joining g_1 to g_2 are homotopic. We may therefore write

$$H = H(g_1, g_2) = H(g(t))$$

where $g(t)$ is any curve joining g_1 to g_2. Rewriting (1), (2) and (3) in this notation, we have

$$H(g_1, g_2) = H(g_0 g_1, g_0 g_2) \tag{4}$$

$$H(g_1, g_3) = H(g_1, g_2)\, H(g_2, g_3) \tag{5}$$

$$H(g_1, g_2) = f(g_1)^{-1} f(g_2) \quad \text{if} \quad g_i \in V. \tag{6}$$

Finally, we define

$$\hat{f}(g) = H(e, g)$$

so that

$$\hat{f}(g) = f(g) \quad \text{for} \quad g \in V \quad \text{by (6)},$$

and

$$\hat{f}(gh) = H(e, gh) = H(e, g)\, H(g, gh) = H(e, g)\, H(e, h) = \hat{f}(g)\hat{f}(h).$$

Thus, $\hat{f}$ is the required homomorphism. Q.E.D.

8. Corollary. If the simply connected groups G_1, G_2 are locally isomorphic, they are isomorphic in the large.

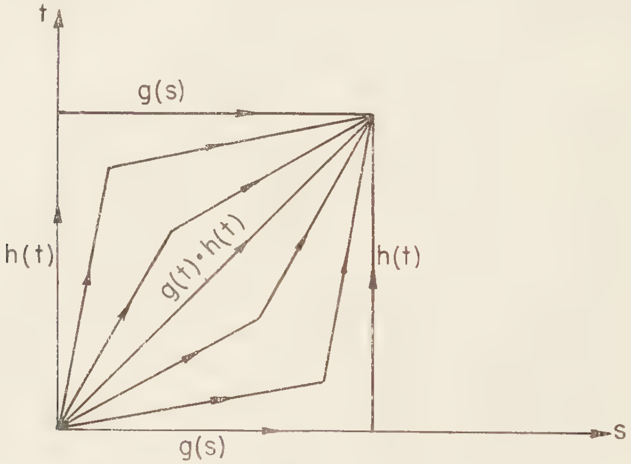
Proof. Let f be the local isomorphism of G_1 into G_2 and let f_1 be its local inverse. We construct $\hat{f}: G_1 \to G_2, \hat{f}_1 : G_2 \to G_1$ as in Theorem 7. We have $f(f_1(g)) = g$ for g near e. Therefore $\hat{f}(\hat{f}_1(g)) = g$ for g near e. But by the uniqueness asserted in Theorem 7, $\hat{f}(\hat{f}_1(g)) = g$ identically. Similarly $\hat{f}_1(\hat{f}(g)) = g$ identically. Q.E.D.

Thus, a simply connected group is completely determined by its local structure. We shall show presently that any Lie group G is locally isomorphic with a simply connected Lie group G, so that there is a $1 - 1$ correspondence between the family of all classes of locally isomorphic Lie groups and the family of all simply connected Lie groups.

The following lemma gives useful information on how the fundamental group $\pi_1(G)$ of a Lie group G is related to the group structure of G and how this fundamental group enters into the question of local isomorphism.

9. Lemma. Let G be a Lie group and let $\pi_1(G)$ be its fundamental group with base point e. Then $\pi_1(G)$ is commutative, and multiplication of paths may be defined directly in terms of the group operation: $[g(t)] * [h(t)] = [g(t) \cdot h(t)] = [h(t)] * [g(t)]$.

Proof. If $g(t)$, $h(t)$ are paths beginning and ending at e, we define $k(s, t) = g(s)\, h(t)$.



The diagram shows the homotopies connecting $g(t) * h(t)$ and $h(t) * g(t)$ with each other and with $g(t)\, h(t)$. Q.E.D.

10. Theorem. Let G be a connected Lie group. Then there is a unique simply connected group $\tilde{G}$ which is locally isomorphic to G. The group $\tilde{G}$ covers G with a projection p which is a homomorphism and local isomorphism. The kernal $p^{-1}(e)$ of the homomorphism p is a discrete subgroup of $\tilde{G}$, isomorphic with $\pi_1(G)$. Each element of $p^{-1}(e)$ commutes with each element of $\tilde{G}$.

Proof. The uniqueness of $\tilde{G}$ up to C^∞ isomorphism follows from Corollary 8. We prove existence and other properties of $\tilde{G}$ by making the universal covering manifold $\tilde{G}$ of G into a Lie group. For convenience, we represent the elements of $\tilde{G}$ as equivalence classes $[g]$ of curves $g = g(t)$ $(0 \leqslant t \leqslant 1)$ in G with $g(0) = e$. Two such curves are to be identified if one can be deformed into the other with fixed end points. By the general theory already outlined in section 0, $\tilde{G}$ may be made into a simply connected covering manifold of G, with the covering projection given by $p([g]) = g(1)$. We make $\tilde{G}$ into a group by defining $[g]\,[h] = [k]$ where $k(t) = g(t)\,h(t)$. Clearly, a homotopy of g and one of h multiply to give a homotopy of k, so this definition is legitimate. The multiplication which this definition gives clearly makes $\tilde{G}$ into a group with identity $e = [e]$, where $e(t) = e$, and inverse defined by $[g]^{-1} = [g^{-1}]$ where $g^{-1}(t) = g(t)^{-1}$. It is also clear from our definition that p is a homomorphism: $p([g]\,[h]) = p([g])\,p([h])$. We know from general considerations that p is a local C^∞ homeomorphism. This shows that the group operations in $\tilde{G}$ are C^∞, making $\tilde{G}$ a Lie group, and that p is a local isomorphism. This completes the first part of the proof.

For the rest, note that $p^{-1}(e)$ is a discrete invariant subgroup of G since it is the kernel of the homomorphism p. If now $g_0 \in p^{-1}(e)$, and $g \in G$, then $f(g) = gg_0g^{-1}$ is a continuous mapping from G into $p^{-1}(e)$. But G is connected, and $p^{-1}(e)$ is discrete. Thus we must have $f(g) =$ constant $= f(e)$, and $gg_0g^{-1} = g_0$ or $gg_0 = g_0g$. Q.E.D.

The following converse of Theorem 10 will be used in a later section.

11. Theorem: Let G be a Lie group, and let N be a discrete invariant subgroup of G. Let $G/N = G'$ be the factor group of G by the subgroup N, and let f be the natural map of G into G', i.e., the map which sends each $g \in G$ into its coset gN. Then we may define a topology for G', and a family of coordinate systems (p_α, U_α) defined in a family of open sets U_α covering G', in such a way that G' becomes a Lie group, that f is a C^∞ map of G onto G' mapping open sets into open sets, and that the mapping f is a covering of G.

Proof: Call a subset U' of G' open if $f^{-1}(U')$ is open. This plainly defines a topology for G', and the map f is plainly continuous. If $U \subseteq G$ is open, then $f^{-1}(f(U)) = NU = \bigcup_{g \in N} gU$ is open; hence f maps open sets onto open sets. Let $g' \neq h'$ be elements of G', and let $f(g) = g', f(h) = h'$. Then we may find a neighborhood V_0 of the identity such that $V_0 = V_0^{-1}$ and $h^{-1}gV_0V_0 \cap N = \phi$. Then $f(gV_0) \cap f(hV_0) = \phi$, since if $v_1, v_2 \in V_0$

and $f(gv_1) = f(hv_2)$, then $gv_1v_2^{-1}h^{-1} \in N$, so $h^{-1}gv_1v_2^{-1} \in N$, a contradiction. Therefore g' and h' are contained in the disjoint open sets $f(gV_0)$ and $f(hV_0)$, proving that G' is a Hausdorff space. If $U' \subseteq G'$ is open, then $f^{-1}(U'^{-1}) = (f^{-1}(U'))^{-1}$; this shows that $g' \to g'^{-1}$ is a continuous map of G' into itself. We leave it to the reader to verify that $(g', h') \to g'h'$ is a continuous map of $G' \times G'$ into G'.

If V is a connected neighborhood of the identity such that $V = V^{-1}$ and $VV \cap N = \{e\}$, and $g \in G$, then f is a one-to-one mapping of gV into G'. For, if $f(gv_1) = f(gv_2)$ $(v_1 \in V)$, we have $f(v_2^{-1}v_1) = e$, so $v_2^{-1}v_1 \in N$, and therefore $v_2^{-1}v_1 = e$, i.e., $v_1 = v_2$.

We may now show that f is a covering of G'. Indeed, if $g' \in G'$, $f^{-1}(g') = gN$ is discrete. Also, $f^{-1}(f(gV)) = \bigcup_{n\in N} ngV$, and it is easily seen that f is a homeomorphism from ngV onto $f(gV)$, and that the component of gn_0 in $\bigcup_{n\in N} (ngV) = \bigcup_{n\in N} (gnV)$ is n_0gV since this latter union is a union of disjoint, connected, open sets.

We now make G' into a manifold in the following natural way. The family $\{f(gV)\}$ covers G'; define coordinate maps by the formula

$$p_g(h') = p[g^{-1}f^{-1}(h')] \quad (h' \in f(gV))$$

where p is the coordinate map for V and f^{-1} sends $f(gv)$ onto gv, $v \in V$. We leave to the reader the straightforward verification that this formula makes G' into a manifold, and also into a Lie group.

The manner in which coordinate systems in G' have been defined plainly makes f a C^∞ map, and a local isomorphism of V onto $f(V)$. This concludes the proof. Q.E.D.

The following analysis reduces the study of connected Lie subgroups to local considerations.

12. Definition. Let G be a Lie group and let H be a submanifold of G which is also a Lie group using the operations of G. Then H is called a Lie subgroup of G.

13. Definition. Let G be a Lie group and let L be a submanifold of G containing the identity e. Suppose there is an open neighborhood U of e in L such that $(g, h) \to gh$ is a C^∞ mapping of $U \times U$ into L and such that $g \to g^{-1}$ is C^∞ from U into L. Then L is called *a local subgroup* of G and U is called a *germ* of the local subgroup.

The reader should bear in mind the example (concerning the torus) which was given immediately preceding Definition 1.27. It partially explains the complications involved in the following theorem.

14. Theorem. Let L be a local subgroup of G, with germ U. Suppose further that U is connected. Then there is a unique connected Lie subgroup $H \subset G$ in which U is a neighborhood of e. Moreover, if L' is a local subgroup of G, U' a connected germ of L', then the connected subgroup H' determined by U' is the same as H if and only if $U \cap U'$ is open in U and in U'.

Proof. We first prove that the mapping $g \to g^{-1}$ sends open subsets of U into open subsets of L and that $g \to g^{-1}$ is C^∞ on the set $U \cup U^{-1}$. To see this, we let $\tau : G \to G$ be the inverse mapping defined by $\tau(g) = g^{-1}$, and $\tau' = \tau | U$. τ is clearly a homeomorphism, and therefore τ_* is $1 - 1$ from Λ_g onto $\Lambda_{g^{-1}}$. If we take into account the natural embedding of the tangent space Λ'_g of L at g into the corresponding tangent space Λ_g of G at g (cf. Lemma 1.30) we see that τ'_* is also $1 - 1$. But τ'_* sends Λ'_g into $\Lambda'_{g^{-1}}$, and these spaces have the same dimension (the dimension of the local subgroup L). Thus τ'_* is also onto and the above assertion is an easy consequence of Lemma 1.18c.

In an entirely analogous manner, we may show that the mappings $g \to g_0 g$, and $h \to h g_0$ are local C^∞ homeomorphisms for fixed $g_0 \in U$.

Now let H be the (abstract) subgroup of G which is generated by U. We now assert that for fixed $h_0 \in H$, there is an open subset U' of U which contains e such that the map $g \to h_0 g h_0^{-1}$ is C^∞ from U' into U. This follows by the above results if we write $h_0 = g_1^{\varepsilon_1} \cdots g_n^{\varepsilon_n}$, $g_i \in U$, $\varepsilon_i = \pm 1$ and use an induction on n.

We now make H into a topological space by defining a set $O \subseteq H$ to be open if for any $g \in O$, there is a neighborhood U' of e in U such that $gU' \subseteq O$. We leave to the reader the verification that this definition makes H a Hausdorff space.

Next, H is made into a C^∞ manifold by the following procedure. Let (p, V) be a coordinate system in U such that $e \in V \subseteq U$, $p(e) = 0$ and $p(V) = \{x \mid |x| < 1, x \in R^k\}$ where k is the dimension of U. We take as a coordinate cover for H the family (p_h, V_h) $(h \in H)$, where $V_h = hV$ and $p_h(h') = p(h^{-1}h')$ for $h' \in hV$. The verification that this makes H into a C^∞ manifold is as follows. If $h_1 V$ and $h_2 V$ overlap, the coordinate transformation connecting these two coordinate neighborhoods is

$$f(x) = p_{h_2} p_{h_1}^{-1}(x) = p(h_2^{-1} h_1 p^{-1}(x)).$$

This mapping of R^k into R^k is C^∞ since it is the composition of C^∞ mappings: $x \to p^{-1}(x) \to (h_2^{-1} h_1) p^{-1}(x) \to p(h_2^{-1} h_1 p^{-1}(x))$.

To show that H is a Lie group, we proceed as follows. Let $h_1, h_2 \in H$. First, there exists a neighborhood V' of e in V such that $V'V' \subseteq V$ (V is the neighborhood of e used to define the C^∞ structure on H.) Then by what has already been proven, there exists a neighborhood V'' of e in V' such that $h_2^{-1}V''h_2 \subseteq V'$, and such that the map $g \to h_2^{-1}gh_2$ is a C^∞ map of V'' into V'. We then have

$$h_2^{-1}V''h_2V'' \subseteq V'V' \subseteq V,$$

and therefore

$$(h_1V'')\,h_2V'' \subseteq h_1h_2V.$$

We now express multiplication $h_1'h_2'$ for $h_1' \in h_iV''$ in coordinates, using the above coordinate systems. If x is the coordinate of h_1' (in the coordinate system (p_{h_1}, h_1V'')), and if y is the coordinate system of h_2' in the coordinate system (p_{h_2}, h_2V''), then multiplication is expressed in coordinates by the function $m(x, y)$, where

$$m(x, y) = p\{(h_1h_2)^{-1}\,[h_1p^{-1}(x)]\cdot[h_2p^{-1}(y)]\}$$

$$= p[h_2^{-2}p^{-1}(x)\,h_2)\cdot p^{-1}(y)].$$

Therefore multiplication is C^∞ from $H \times H$ into H. Similarly, the mapping $h \to h^{-1}$ is C^∞ from H into H, and therefore H is a Lie group.

To show that H is a Lie sub-group of G, it is necessary to show that the inclusion map $i: H \to G$ is regular (i.e., that i_* is $1-1$). But this is trivial at e and hence at all $h \in H$ by translation.

Finally, H is connected by the following argument. The component H_0 of e in H is certainly connected (cf. Lemma 3). Clearly U, by hypothesis connected, is in H_0, and is a neighborhood of e in H_0. Therefore, by Lemma 2, H_0 is generated, group-theoretically, by U. But, by definition, H is generated by U. Hence $H = H_0$.

We have thus far shown the existence of the connected Lie subgroup H described in the statement of the theorem. The uniqueness is clear; if H_1 satisfied the hypotheses of the theorem, it would of course have to be identical with H as far as its points, its topology and its C^∞ structure.

Finally, suppose the connected germ U' generates the connected Lie subgroup H', with U' open in H'. Then if $H = H'$, U and U' are both open in H, hence $U \cap U'$ is open in U and U'. Now suppose conversely $U \cap U'$ is open in U. It then follows that $U \cap U'$ generates H, since $U \cap U'$ is a neighborhood of e which is contained in a connected neighborhood of e. Therefore, if $U \cap U'$ is open in U and U', it follows that $U \cap U'$ generates both H and H' and therefore $H = H'$, verifying the final statement of our lemma. Q.E.D.

3. The Lie Algebra

In this section and the following sections, we show how the study of a Lie group G may be greatly simplified by considering the tangent space Λ of G at e. We shall show how a *multiplication* (the Lie product, or Lie bracket) may be introduced in Λ and that the resulting algebraic structure—called a Lie algebra—determines the local structure of the group. Thus, two groups will be locally isomorphic if and only if their Lie algebras are isomorphic. The Lie algebra is a finite dimensional algebra. Therefore the local study of Lie group is entirely equivalent to the study of certain finite dimensional linear algebraic structures—a great simplification.

If $g \in G$, let Λ_g denote the tangent space of G at $g \in G$. Then the left multiplication $h \to g_0 h$ induces (cf. Definition 1.14) a linear map from Λ_h into $\Lambda_{g_0 h}$. Similarly, right multiplication by g_0 induces a map of Λ_h into Λ_{hg_0}.

1. Definition. If $\lambda \in \Lambda_h$, we write $g\lambda$ for the image of λ under the map of Λ_h into Λ_{gh} induced by the map $h \to gh$, and λg for the image of λ under the map of Λ_h into Λ_{hg} induced by the map $h \to hg$.

Using Lemma 1.21, this definition immediately implies the following lemma.

2. Lemma. Let $\lambda \in \Lambda_h$ be represented by the curve $h(t)$ at $t = t_0$, and let $g \in G$. Then $g\lambda$ is the element of Λ_{gh} represented by the curve $g \cdot h(t)$ at $t = t_0$, and $\lambda g \in \Lambda_{hg}$ is the tangent vector represented by $h(t)\, g$ at $t = t_0$.

In symbols, (cf. Definition 1.19):

$$\frac{d}{dt}\, gh(t) = g\, \frac{dh(t)}{dt}\, ; \quad \frac{d}{dt}\, h(t)\, g = \frac{dh(t)}{dt}\, g \tag{1}$$

By applying Lemma 1.15 to the operations of left and right translation, we easily obtain the following properties of the multiplication introduced in Definition 1.

3. Lemma. Let $g, h, k \in G$, $\lambda, \mu \in \Lambda_h$, a, b real. Then

$$g(a\lambda + b\mu) = a(g\lambda) + b(g\mu)$$

$$(a\lambda + b\mu)g = a(\lambda g) + b(\mu g)$$

$$g(k\lambda) = (gk)\,\lambda; \quad (\lambda g)\,k = \lambda(gk)$$

$$(g\lambda)\,k = g(\lambda k)\,(= g\lambda k)$$

$$e\lambda = \lambda e = \lambda$$

Formula (1) may be generalized to obtain the derivative of $g(t)\,h(t)$. The usual formulas of calculus apply.

4. Lemma. If $g(t)$, $h(t)$ are curves in G, then

$$\frac{d}{dt}\,g(t)\,h(t) = g(t)\,\frac{dh}{dt} + \frac{dg}{dt}\,h(t) \tag{2}$$

$$\frac{d}{dt}\,[g(t)]^{-1} = -g^{-1}\,\frac{dg}{dt}\,g^{-1} \tag{3}$$

Proof. We calculate $\dfrac{d}{dt}\,g(t)\,h(t)$ by using the definition (Definition 1.19) and the chain rule. Let f be a real-valued function on G. Then

$$\frac{d}{dt}\,[g(t)\,h(t)]\Big|_{t=t_0}(f) = \frac{d}{dt}\,f(g(t)\,h(t))\Big|_{t=t_0}$$

$$= \frac{\partial}{\partial s}\,f(g(s)\,h(u))\,\frac{ds}{dt}\Big|_{t=t_0} + \frac{\partial}{\partial u}\,f(g(s)\,h(u))\,\frac{du}{dt}\Big|_{t=t_0}$$

where $s = s(t) = t$ and $u = u(t) = t$.

Thus,

$$\frac{d}{dt}\,[g(t)\,h(t)]\Big|_{t=t_0}(f) = \frac{d}{dt}\,f(g(t)\,h(t_0))\Big|_{t=t_0} + \frac{d}{dt}\,f(g(t_0)\,h(t))\Big|_{t=t_0}$$

$$= \frac{dg}{dt}\Big|_{t=t_0}\,h(t_0)\,(f) + g(t_0)\,\frac{dh}{dt}\Big|_{t=t_0}(f)$$

$$= \left[g(t_0)\,\frac{dh}{dt}\Big|_{t=t_0} + \frac{dg}{dt}\Big|_{t=t_0}\,h(t_0)\right](f)$$

This proves (2). Formula (3) follows by a straight-forward use of the calculus of implicit functions:

$$0 = \frac{de}{dt} = \frac{d}{dt}(g(t)\,[g(t)]^{-1}) = g\frac{d}{dt}(g^{-1}) + \frac{dg}{dt}g^{-1}$$

Solving for $\dfrac{d}{dt}g^{-1}$, with the help of Lemma 3, we have (3). Q.E.D.

Since $g\lambda$ and λg are defined in terms of the group structure, the following lemma is clear.

5. Lemma. Let $f: G \to G'$ be a homomorphism, $g, h \in G$, $g' = f(g)$, $h' = f(h)$, so that $f_*: \Lambda_h \to \Lambda_{h'}, f_*: \Lambda_{gh} \to \Lambda_{g'h'}, f_*: \Lambda_{hg} \to \Lambda_{h'g'}$. Then

$$f_*(\lambda g) = f_*(\lambda) f(g)$$

$$f_*(g\lambda) = f(g) f_*(\lambda)$$

for $\lambda \in \Lambda_h$.

Proof. Let λ be represented by the curve $h(t)$ at $t = 0$, so that $dh/dt|_{t=0} = \lambda$. Then, using Lemma 1.21 and the homomorphism property of f, we have

$$f_*(\lambda g) = f_*\left(\left.\frac{dh}{dt}\right|_{t=0} g\right) = f_*\left(\left.\frac{d(h(t)\,g)}{dt}\right|_{t=0}\right) = \left.\frac{d}{dt} f(h(t)) f(g)\right|_{t=0}$$

$$= \left.\frac{d}{dt} f(h(t))\right|_{t=0} f(g) = f_*\left(\left.\frac{dh}{dt}\right|_{t=0}\right) f(g) = f_*(\lambda) f(g)$$

A similar argument is used for $f_*(g\lambda)$. Q.E.D.

Remark. If f is only a local homomorphism, then the above proof shows that, for example, there exists a neighborhood U of e in G such that if $g, h \in U$, $\lambda \in \Lambda_h$, then $f_*(\lambda g) = f_*(\lambda) f(g)$.

Now let $\Lambda = \Lambda_e$ be the tangent space of G at the identity e. Lemma 4 can be used to show how *the group structure of* G and the *vector space structure* of Λ are related.

6. Lemma. Let $\lambda, \mu \in \Lambda$ and let c be real. Suppose λ is represented by $g(t)$, μ by $h(t)$ at $t = 0$. Then

$\lambda + \mu$ is represented by $g(t)\,h(t)$
$-\lambda$ is represented by $[g(t)]^{-1}$
$c\lambda$ is represented by $g(ct)$

Proof. By definition, we have

$$\lambda = \left.\frac{dg}{dt}\right|_{t=0}, \ \mu = \left.\frac{dh}{dt}\right|_{t=0}, \ g(0) = h(0) = e.$$

Thus,

$$\left.\frac{d}{dt}\, gh\right|_{t=0} = g(0)\left.\frac{dh}{dt}\right|_{t=0} + \left.\frac{dg}{dt}\right|_{t=0} h(0) = \lambda + \mu$$

$$\left.\frac{d}{dt}\, [g(t)]^{-1}\right|_{t=0} = -g(0)^{-1}\left.\frac{dg}{dt}\right|_{t=0} g(0) = -\lambda$$

$$\left.\frac{d}{dt}\, g(ct)\right|_{t=0} = c\left.\frac{dg}{dt}\right|_{t=0} = c\lambda$$

These are our assertions. Q.E.D.

We now take a basic step by introducing a *multiplication* in Λ.

7. Definition (of Lie bracket): Let $\lambda, \mu \in \Lambda$ and let λ be represented by the curve $g(t)$ at $t = 0$. We define

$$[\lambda, \mu] = \left.\frac{d}{dt}\,(g(t)\,\mu[g(t)]^{-1})\right|_{t=0}$$

More explicitly, if $g(t)$ is any curve satisfying $g(0) = e$, we have

$$\left[\left.\frac{dg}{dt}\right|_{t=0}, \mu\right] = \left.\frac{d}{dt}\,(g(t)\,\mu[g(t)]^{-1})\right|_{t=0}$$

Using Lemma 1.21, we can show very easily that $[\lambda, \mu]$ does not depend upon the choice of the curve $g(t)$ used to represent λ. To see this, we take $M = G$ and $M' = \Lambda$. If we fix $\mu \in \Lambda$ and define the function $f\colon M \to M'$ by the formula $f(g) = g\mu g^{-1}$, it is seen that Definition 7 merely states that $[\lambda, \mu] = f_*(\lambda)$, where a tangent vector to the vector space Λ identified with a vector of Λ in the usual way.

Note that $[\lambda, \mu] \in \Lambda$ if $\lambda, \mu \in \Lambda$, so that $[\lambda, \mu]$ may be regarded as a product of elements of Λ, with a value in Λ.

In order to compute $[\lambda, \mu]$, and reveal its properties, we go to coordinates. We take a coordinate system (p, U) containing e, and assume that $p(e) = 0$. We assume G is n-dimensional and express multiplication by a function $z = m(x, y)$ where $x, y, z \in R^n$. Explicitly

$$m(x, y) = p[p^{-1}(x) \cdot p^{-1}(y)]$$

where x and y lie in a sufficiently small neighborhood of 0. Similarly, the inverse function $w = I(x)$ is expressed by

$$I(x) = p([p^{-1}(x)]^{-1})$$

From $ge = eg = g$, it follows that

$$m(x, 0) = x, \qquad m(0, y) = y \tag{4}$$

Setting

$$m(x, y) = a(x) + b(y) + c(x, x) + d(x, y) + e(y, y) + \cdots$$

and using $m(0, 0) = 0$, we have

$$x = a(x) + c(x, x) + \cdots; \qquad y = b(y) + e(y, y) + \cdots$$

by (4). By comparing both sides, we find $a(x) = x$, $c(x, x) = 0$, $b(y) = y$, $e(y, y) = 0, \ldots$ so that (changing notation slightly),

$$m(x, y) = x + y + a(x, y) + \cdots \tag{5}$$

The same argument gives

$$m(x, y) = x + y + a(x, y) + b(x, x, y) + c(x, y, y) + \cdots \tag{6}$$

to third order. At this point, we note that the second order bilinear term $a(x, y)$ depends on the coordinate system, yet the first order term is always $x + y$ in any coordinate system. Thus we may say that multiplication is commutative to first order.

Next we calculate the form of the inverse function $I(x)$. Since $gg^{-1} = e$, we have, in coordinates, $m(x, I(x)) = 0$. Hence, if

$$I(x) = d(x) + e(x, x) + \cdots$$

to second order, then substituting in (5), we have

$$\begin{aligned} 0 = m(x, I(x)) &= x + I(x) + a(x, I(x)) + \cdots \\ &= x + d(x) + e(x, x) + a(x, d(x)) + \cdots \end{aligned}$$

By comparing terms, we have $x + d(x) = 0$; $e(x, x) + a(x, d(x)) = 0$ Hence $d(x) = -x$, $e(x, x) = a(x, x)$, and

$$I(x) = -x + a(x, x) + \cdots \tag{7}$$

The following lemma gives the coordinate expressions for various important group-theoretical expressions.

8. Lemma. Let x and y be the coordinates of g and h in a coordinate system about e. Let $\tilde{\lambda}$ denote the coordinate of the tangent vector λ in this coordinate system. Let "$\sim$" denote "is expressible by the function" so that

$$gh \sim m(x, y) = x + y + a(x, y) + \cdots \tag{5}$$

$$g^{-1} \sim I(x) = -x + a(x, x) + \cdots \tag{7}$$

Then

$$ghg^{-1} \sim y + a(x, y) - a(y, x) + \cdots \tag{8}$$

and

$$ghg^{-1}h^{-1} \sim a(x, y) - a(y, x) + \cdots \tag{9}$$

to second order in x and y. Moreover

$$\widetilde{g\lambda h} = \tilde{\lambda} + a(x, \tilde{\lambda}) + a(\tilde{\lambda}, y) + \cdots \tag{10}$$

to first order in x and y for $\lambda \in \Lambda$, and

$$\widetilde{g\mu g^{-1}} = \tilde{\mu} + a(x, \tilde{\mu}) - a(\tilde{\mu}, x) + \cdots \tag{11}$$

to first order in x for $\mu \in \Lambda$. Finally

$$\widetilde{[\lambda, \mu]} = a(\tilde{\lambda}, \tilde{\mu}) - a(\tilde{\mu}, \tilde{\lambda}). \tag{12}$$

Proof. By computation we have

$$\begin{aligned} ghg^{-1} &\sim m(m(x, y),\ I(x)) = m(x, y) + I(x) + a(m(x, y),\ I(x)) + \cdots \\ &= x + y + a(x, y) + -x + a(x, x) + a(x + y, -x) + \cdots \\ &= y + a(x, y) - a(y, x) + \cdots \end{aligned}$$

This proves (8). Using (8) we have

$$\begin{aligned} ghg^{-1}h^{-1} &\sim m(y + a(x, y) - a(y, x) + \cdots, I(y)) \\ &= y + a(x, y) - a(y, x) + -y + a(y, y) + a(y, -y) + \cdots \\ &= a(x, y) - a(y, x) + \cdots \end{aligned}$$

which proves (9). To prove (10), we suppose that λ is represented by the curve $k(t)$ and

$$k(t) \sim \tilde{\lambda}t$$

Then $g\lambda h$ is represented by $gk(t)\,h$. But

$$gk(t) \sim x + \tilde{\lambda}t + a(x, \tilde{\lambda}t) + \cdots = x + (\tilde{\lambda} + a(x, \tilde{\lambda}) + \cdots)\, t + \cdots$$

Hence

$$\begin{aligned} gk(t)\,h \sim{} & x + (\tilde{\lambda} + a(x, \tilde{\lambda}) + \cdots)\, t + y \\ & + a(x + (\tilde{\lambda} + a(x, \tilde{\lambda}) + \cdots)\, t, y) + \cdots \\ & = x + y + a(x, y) + \cdots + (\tilde{\lambda} + a\,(x,\tilde{\lambda}) \\ & + a(\tilde{\lambda}, y) + \cdots)\, t + \cdots \end{aligned}$$

Thus, using Lemma 1.12, we have (10). (11) may be proved in a more direct manner from either (8) or (10). Finally, if $g(t)$ represents λ at $t = 0$, we have

$$g(t) \sim \tilde{\lambda}t + \cdots$$

and by (11) we have

$$\widetilde{g(t)\,\mu[g(t)]^{-1}} = \tilde{\mu} + [a(\tilde{\lambda}, \tilde{\mu}) - a(\tilde{\mu}, \tilde{\lambda})]\, t + \cdots$$

Hence we have (12) by Definition 7. Q.E.D.

Infinitesimal statements of the formulae of Lemma 8 are fairly clear and suggestive. Informally, they are:

$$(e + \lambda dt)\,(e + \mu ds)\,(e + \lambda dt)^{-1} = e + \mu ds + [\lambda, \mu]\, dsdt \tag{8$'$}$$

$$(e + \lambda dt)\,(e + \mu ds)\,(e + \lambda dt)^{-1}\,(e + \mu ds)^{-1} = e + [\lambda, \mu]\, dsdt \tag{9$'$}$$

$$(e + \lambda dt)\,\mu(e + \lambda dt)^{-1} = \mu + [\lambda, \mu]\, dt \tag{11$'$}$$

Formulas (6) and (7) give

$$(e + \lambda dt)\,(e + \mu dt) = e + (\lambda + \mu)\, dt \tag{6$'$}$$

$$(e + \lambda\, dt)^{-1} = e - \lambda\, dt. \tag{7$'$}$$

We now prove some general properties about the Lie bracket.

9. Theorem. Let $\lambda, \mu, \nu \in \Lambda$, and let a and b be real. Then

$$[\lambda, a\mu + b\nu] = a[\lambda, \mu] + b[\lambda, \nu]$$

$$[a\mu + b\nu, \lambda] = a[\mu, \lambda] + b[\nu, \lambda]$$

Thus, $[\lambda, \mu]$ is a bilinear function from $\Lambda \times \Lambda$ into Λ. Moreover

$$[\lambda, \mu] = -[\mu, \lambda] \quad \text{and} \quad [\lambda, \lambda] = 0. \quad \text{(skew-symmetry)} \tag{13}$$

$$[[\lambda, \mu], \nu] + [[\mu, \nu], \lambda] + [[\nu, \lambda], \mu] = 0. \quad \text{(Jacobi identity)} \tag{14}$$

Proof. We work entirely in a coordinate system. Bilinearity of $[\lambda, \mu]$ follows from the formula

$$\widetilde{[\lambda, \mu]} = a(\tilde{\lambda}, \tilde{\mu}) - a(\tilde{\mu}, \tilde{\lambda}) \tag{12}$$

of Lemma 8. Formula (13) is an immediate consequence of formula (12), for this implies $[\lambda, \mu] = -[\mu, \lambda]$ and $[\lambda, \lambda] = 0$. (It is an easy consequence of bilinearity that either of the equalities in (13) implies the other.)

The proof of (14) involves a lengthier computation and is the first full application of the associative law in G. If g, h, k have coordinates x, y, z, we have (using the "$\sim$" symbol of Lemma 8),

$$(gh)\,k \sim m(m(x, y), z); \quad g(hk) \sim m(x, m(y, z))$$

Thus,

$$m(m(x, y), z) = m(x, m(y, z)). \tag{15}$$

We calculate to the third order, as (15) turns out to be true automatically to the second order. We have, by the computations preceding Lemma 8

$$m(x, y) = x + y + a(x, y) + b(x, x, y) + c(x, y, y) + \cdots \tag{6}$$

Thus, by a straight-forward computation,

$$m(m(x, y), z) = \cdots + a(a(x, y), z) + b(x, y, z) + b(y, x, z) + \cdots$$

where we have omitted second order terms and all third order terms with a "repeated factor". We eliminate the terms involving b by interchanging x and y and subtracting, to obtain

$$m(m(x, y), z) - m(m(y, x), z) = \cdots + a([x, y], z) + \cdots \tag{16}$$

where we have adopted the natural notation $[x, y] = a(x, y) - a(y, x)$. In exactly the same way, we obtain

$$m(z, m(x, y)) - m(z, m(y, x)) = \cdots + a(z, [x, y]) + \cdots \tag{17}$$

We now subtract (17) from (16) and use the more suggestive $x * y$ instead of $m(x, y)$, to obtain

$$(x * y) * z - z * (x * y) + z * (y * x) - (y * x) * z$$

$$= \cdots + [[x, y], z] + \cdots \tag{18}$$

Next we rewrite (18) after changing x to y, y to z and z to x.

$$(y*z)*x - x*(y*z) + x*(z*y) - (z*y)*x = \cdots + [[y, z], x] + \cdots \tag{19}$$

and similarly

$$(z * x) * y - y * (z * x) + y * (x * z) - (x * z) * y$$

$$= \cdots + [[z, x], y] + \cdots \tag{20}$$

Now add (18) through (20) and use the associative law (15), i.e. $(x * y) * z = x * (y * z)$, obtaining

$$0 = \cdots + [[x, y], z] + [[y, z], x] + [[z, x], y] + \cdots \tag{21}$$

Finally, by comparing like terms and using Lemma 8, we have Jacobi's identity (14). Q.E.D.

Theorem 9 suggests the following definition.

10. Definition. Let V be a (finite dimensional) vector space in which a function $(\lambda, \mu) \to \lambda\mu$ from $V \times V$ into V is defined. If $\lambda\mu$ is bilinear as a function of λ and μ, V is called an *algebra*. If further,

$$\lambda\lambda = 0$$

and

$$(\lambda\mu)\,\nu + (\mu\nu)\,\lambda + (\nu\lambda)\,\mu = 0$$

V is called a *Lie algebra*, and we write $\lambda\mu = [\lambda, \mu]$.

The condition $\lambda\lambda = 0$ implies skew symmetry:

$$\lambda\mu = -\mu\lambda, \text{ since } \lambda\mu + \mu\lambda = (\lambda + \mu)(\lambda + \mu) - \lambda\lambda - \mu\mu = 0.$$

Theorem 9 states that Λ is a Lie algebra in the sense of Definition 10.

11. Definition. Let Λ_1 and Λ_2 be Lie algebras, and let $\sigma\colon \Lambda_1 \to \Lambda_2$. σ is called a *homomorphism* if it is linear and $\sigma[\lambda, \mu] = [\sigma(\lambda), \sigma(\mu)]$. If σ is also $1 - 1$ and onto, σ is called an *isomorphism*. (In this latter case it follows, of course, that σ^{-1} is also an isomorphism.)

We note that the Lie algebra of a Lie group is definable in terms of the local structure of the group. Indeed, the tangent space at e is so definable and the definition of $[\lambda, \mu]$ and its properties involve the multiplication in the group only locally. For these reasons we will generally frame theorems connecting Lie groups and their Lie algebras locally.

12. Lemma. Let $f: G \to G'$ be a local homomorphism of G into G', and let Λ, Λ' be the Lie algebras of G, G' respectively. Then f_* is a homomorphism of Λ into Λ'. If f is a local isomorphism, then f_* is an isomorphism.

Proof. To prove our first assertion, it is only necessary to prove that f_* preserves Lie products. Let $\lambda, \mu \in \Lambda$ and suppose λ is represented by $g(t)$. Then, by Definition 6,

$$[\lambda, \mu] = \frac{d}{dt}\, g(t) \cdot \mu g(t)^{-1}\Big|_{t=0}$$

Hence, using Lemma 5, we have

$$f_*[\lambda, \mu] = f_* \frac{d}{dt}\, g(t)\, \mu g(t)^{-1}\Big|_{t=0} = \frac{d}{dt}\, f_*\, g(t) \mu g(t)^{-1}\Big|_{t=0}$$

$$= \frac{d}{dt} fg(t)\, f_*\mu [fg(t)]^{-1}\Big|_{t=0} = [f_*\lambda, f_*\mu]$$

since $fg(t)$ represents $f_*\lambda$. This proves the first assertion; the second follows from the first and from Lemma 1.15. Q.E.D.

13. Corollary. If $g \in G$ and $\lambda, \mu \in \Lambda$, then

$$g[\lambda, \mu]\, g^{-1} = [g\lambda g^{-1}, g\mu g^{-1}]$$

Proof. For fixed g, the map $h \to ghg^{-1}$ is an automorphism τ of G into G for which $\tau_*\lambda = g\lambda g^{-1}$. Q.E.D.

If $g \in G$ and if $\lambda \in \Lambda_g$, then $g^{-1}\lambda \in \Lambda_e$. This process of "translation" gives a natural way of comparing tangent vectors at different points of a Lie group with tangent vectors at the identity, and hence with each other, and hence, ultimately, of differentiating variable families of vectors, etc. This is the circumstance which enables us to state the following lemma which generalizes the "integrability conditions" for a gradient vector field to an arbitrary Lie group.

14. Lemma. Let $g = g(s, t)$ be a function of two real variables into G and let

$$\alpha = \alpha(s, t) = g^{-1} \frac{\partial g}{\partial s}\,; \beta = g^{-1} \frac{\partial g}{\partial t}$$

Then $\alpha, \beta \in \Lambda$ and we have

$$\frac{\partial\beta}{\partial s} - \frac{\partial\alpha}{\partial t} + [\alpha, \beta] = 0.$$

Proof. We have only to verify this equation at a particular point (s, t), say $(s, t) = (0, 0)$. We next note that we may assume $g(0, 0) = e$, for we may replace g by $h = g^{-1}(0, 0)\, g(s, t)$ and prove the result for h and hence for the same α and β. We now go to coordinates where multiplication is given by

$$m(x, y) = x + y + a(x, y) + \cdots$$

We suppose $g = g(s, t)$ is given in coordinates by

$$\tilde{g} = x(s, t) = sb + tc + s^2e + 2stf + t^2h + \cdots$$

(where b, c, e, f g are vectors in R^n). Then

$$\frac{\widetilde{\partial g}}{\partial s} = b + 2se + 2tf + \cdots$$

$$\frac{\widetilde{\partial g}}{\partial t} = c + 2sf + 2th + \cdots$$

Since $g\alpha = \dfrac{\partial g}{\partial s}$, $g\beta = \dfrac{\partial g}{\partial t}$, we may use Lemma 8 (equation (10)) to obtain

$$\tilde{\alpha} + a(sb + tc + \cdots, \tilde{\alpha}) = b + 2se + 2tf + \cdots$$

$$\tilde{\beta} + a(sb + tc + \cdots, \tilde{\beta}) = c + 2sf + 2th + \cdots$$

These equations give $\tilde{\alpha}$ and $\tilde{\beta}$ to first order (by comparing coefficients):

$$\tilde{\alpha} = b + 2se + 2tf - a(sb + tc, b) + \cdots$$

$$\tilde{\beta} = c + 2sf + 2th - a(sb + tc, c) + \cdots$$

Hence $\left.\dfrac{\widetilde{\partial\alpha}}{\partial t}\right|_{(0,0)} = 2f - a(c, b)$

$$\left.\frac{\widetilde{\partial\beta}}{\partial s}\right|_{(0,0)} = 2f - a(b, c)$$

But by formula (12) of Lemma 8, we have

$$\widetilde{[\alpha, \beta]}\Big|_{(0,0)} = a(b, c) - a(c, b),$$

giving the result. Q.E.D.

Matrices and linear transformations play a central role in the study of Lie groups and Lie algebras. We therefore now briefly describe and consider the particular Lie group of all non-singular $n \times n$ matrices.

We let $A(n)$ be the set of all real $n \times n$ matrices. Then $A(n)$ is an n^2-dimensional vector space, and it is therefore also a C^∞ manifold. It is also an (associative) algebra, since it is endowed with a multiplication (matrix multiplication). The determinant of S (written $\det S$) is clearly a real-valued C^∞ function of S since it is a polynomial in the entries of S. Since the non-singular matrices are characterized by the equation $\det S \neq 0$, such matrices form an open set in $A(n)$, and therefore have a C^∞ structure induced by $A(n)$. The *general linear group of order n, denoted GL (n), is defined as the group of all non-singular $n \times n$ matrices.* It is clearly a group, and in fact a Lie group of dimension n^2. Indeed, multiplication has already been shown to be C^∞ and by "Cramer's Rule" the inverse operation $S \to S^{-1}$ involves only polynomial operations and division by a non-zero determinant. $GL(n)$ is not connected since the values of $\det S$ are all non-zero numbers, a disconnected set.

A similar analysis shows that if we permit complex entries in the matrices of $GL(n)$, we also obtain a Lie group. We then use the notation $GL(n, C)$ and the (real) dimension is $2n^2$. However, we shall now confine our detailed comments to the real matrices and merely remark that the analysis of $GL(n, C)$ is entirely similar.

The identity e of $GL(n)$ is the identity matrix I. If G is an open set of a vector space V, the identity mapping: $\text{Ident}: G \to V$ gives a coordinate system covering all of G, and therefore, using this coordinate system, the tangent space to G at a point $P \in G$ is identified with V itself in a natural way. Thus, the tangent space Λ of $GL(n)$ may be taken to be $A(n)$—the linear space of all matrices. Then, for example, a matrix $\lambda \in A(n)$ is represented (in the sense of Definition 1.19) by the curve $S(t) = I + t\lambda$. More generally, any curve $S'(t)$ with $S(0) = I$ represents the tangent vector (matrix) $\lambda = \dfrac{dS'}{dt}\Big|_{t=0}$.

In the same way, $A(n)$ is also identified with the tangent space at any $S_0 \in GL(n)$. Definition 1 of this section is entirely consistent with standard matrix notation. For a tangent vector at $S \in GL(n)$ is given by a matrix λ

and is represented by the curve $S + t\lambda$ at $t = 0$. The left-translate of this curve by T is the curve $T(S + t\lambda) = TS + tT\lambda$, which by Lemma 2 represents $T\lambda$ in the sense of Definition 3.1. Thus the tangent vector $T\lambda$ (and similarly λT) of Definition 3.1 is obtained by matrix multiplication.

We may now compute $[\lambda, \mu]$ for $\lambda, \mu \in A(n)$ (the Lie algebra of $GL(n)$). We let $S(t)$ represent λ, so that $S(0) = I$ and $\acute{S}(0) = \lambda$. Then by using the ordinary calculus of matrices and Definition 7, we have

$$[\lambda, \mu] = \frac{d}{dt}(S\mu S^{-1})\Big|_{t=0} = \frac{dS}{dt}\,\mu S^{-1}\Big|_{t=0} + S\mu \cdot -S^{-1}\frac{dS}{dt}\,S^{-1}\Big|_{t=0}$$

$$= \lambda\mu - \mu\lambda.$$

We remark that if V is any (finite dimensional) vector space, and if $E(V)$ is the algebra of linear mappings of V into V, and if $GL(V)$ is the group of non-singular linear mappings of V into V, then the above analysis shows that $GL(V)$ is in a natural way a Lie group whose Lie algebras is $E(V)$. The Lie bracket in $E(V)$ is given by the formula $[\sigma, \tau] = \sigma\tau - \tau\sigma$.

We summarize the last few paragraphs in the following Lemma.

15. Lemma. Let $GL(n)$ be the Lie group of all non-singular $n \times n$ matrices, and let $A(n)$ be the algebra of all $n \times n$ matrices. Then $A(n)$ is isomorphic with, and identified with the Lie algebra of $GL(n)$, where the Lie bracket operation in $A(n)$ is given by the formula $[S, T] = ST - TS$. Under this identification, the tangent vector represented by a curve $S(t)$ in $GL(n)$ with $S(0) = I$ is given by the matrix $\lambda = (dS/dt)|_{t=0}$; a similar formula identifies the tangent vector at any matrix of $GL(n)$ with a matrix of $A(n)$. The translation operations $\lambda \to g\lambda$ and $\lambda \to \lambda g$ is given by matrix multiplication of the matrix g and the matrix λ.

4. The Exponential Function

In a discrete group, any element g generates a subgroup, consisting of all powers of g, which is cyclic and abelian. For Lie groups, infinitesimal elements ($\lambda \in \Lambda$) generate a subgroup which is abelian and parametrized by a real parameter t. The "infinitesimal generation" is suggested as follows. If $g(t)$ is generated by $(e + \lambda\,dt)$, we mean $g(t + dt) = g(t)\,(e + \lambda\,dt)$. (Compare with $g^{n+1} = g^n g$.) Thus, $g + dg = g + g\lambda dt$ and $dg/dt = g\lambda$, and we have the exponential.

1. Lemma. Let $\lambda \in \Lambda$. Then the equation

$$\frac{dg}{dt} = g\lambda, \quad g(0) = e \tag{1}$$

has a unique solution $g = g(t) \in G$ for $-\infty < t < +\infty$.

Proof. By Lemma 1.24, equation (1) has a local solution near $t = 0$, defined for $0 \leqslant t \leqslant a$ with $a > 0$. Call this solution $g_1(t)$. We extend $g_1(t)$ to all $t \geqslant 0$ by induction. Let $g_n(t)$ be defined for $0 \leqslant t \leqslant na$. Then define $g_{n+1}(t)$ by the formulas:

$$\begin{aligned} g_{n+1}(t) &= g_n(t) && 0 \leqslant t \leqslant na \\ &= g_n(na)\, g_1(t - na) && na \leqslant t \leqslant (n+1)\,a \end{aligned}$$

so that $g_{n+1}(t)$ is an extension of $g_n(t)$. It is trivial by induction that $g_{n+1}(t)$ satisfies (1) for $t \in [0, (n+1)\,a]$. In this way, we construct a function $g^+(t) = g_\lambda^+(t)$ which satisfies (1) for $t \geqslant 0$:

$$\frac{dg_\lambda^+(t)}{dt} = g_\lambda^+(t)\,\lambda, \quad 0 \leqslant t < \infty, g(0) = e.$$

Finally, we define

$$g_\lambda^-(t) = g_{-\lambda}^+(-t) \qquad -\infty < t \leqslant 0$$

and we find that

$$\frac{d}{dt} g_\lambda^-(t) = \frac{d}{dt} g_{-\lambda}^+(-t) = -1 \frac{d}{ds} g_{-\lambda}^+(s)\Big|_{s=-t}$$

$$= -g_{-\lambda}^+(-t)\,(-\lambda) = g_\lambda^-(t)\,\lambda$$

so that $g_\lambda^-(t)$ satisfies (1) for $-\infty < t \leqslant 0$. Since $g_\lambda^-(0) = g_\lambda^+(0)$, we have constructed $g_\lambda(t)$ for all $-\infty < t < \infty$.

Uniqueness is trivial but we prove it now and omit such proofs in similar arguments. If $g_1(t)$, $g_2(t)$ both satisfy (1), then let $S =$ the set of values of t where $g_1(t) = g_2(t)$. Then S is closed since $g_1(t)$ and $g_2(t)$ are continuous, while S is open because of the local uniqueness of solutions of (1). Since $0 \in S$, S is non-empty. Hence S is the set of all real numbers, since the real numbers are connected. Q.E.D.

Remark. By Lemma 1.25, $g_\lambda(t)$ is C^∞ as a function of (λ, t).

2. Definition. Let $g_\lambda(t)$ be the solution of equation (1) of Lemma 1. Then we put $\exp(\lambda) = g_\lambda(1)$. We also write $\exp(\lambda) = e^\lambda$.

We remark that while $\lambda \in \Lambda$, $e^{\lambda} \in G$. Thus, the exponential function brings us from the Lie algebra into the Lie group.

3. Lemma. The solution of equation (1) is $g(t) = \exp(t\lambda)$. Thus,

$$\frac{d}{dt} \exp(t\lambda) = [\exp(t\lambda)]\,\lambda$$

Proof. Let $g_{\lambda}(t)$ be the solution of (1) and let s be an arbitrary real number. Let $h(t) = g_{\lambda}(st)$. Then

$$\frac{dh}{dt} = s \left.\frac{dg_{\lambda}(r)}{dr}\right|_{r=st} = sg_{\lambda}(st)\,\lambda = h(t)\,(s\lambda)$$

Moreover $h(0) = e$. If we compare this latter differential equation with (1). we have

$$g_{s\lambda}(t) = h(t) = g_{\lambda}(st) \tag{2}$$

Putting $t = 1$ in (2) gives $\exp(s\lambda) = g_{\lambda}(s)$, which is the result. Q.E.D.

4: Theorem. For fixed $\lambda \in \Lambda$, the mapping $t \to \exp(t\lambda)$ is a homomorphism of the additive group of reals into G. Thus,

$$\exp(s + t)\,\lambda = \exp s\,\lambda \exp t\lambda.$$

Proof: Let s_0 be fixed and let $h(t) = (\exp s_0\lambda)^{-1}.\ \exp[(s_0 + t)\,\lambda]$. Then $h(0) = e$ and it follows by Lemma 3 that $\dfrac{dh}{dt} = h\lambda$. Therefore $h(t)$ satisfies (1). Using Lemma 3 again, it follows that $h(t) = \exp t\lambda$. Hence

$$\exp t\lambda = (\exp s_0\lambda)^{-1} \exp(s_0 + t)\,\lambda$$

which is the desired result. Q.E.D.

5. Corollary. $\exp t\lambda$ and $\exp s\lambda$ commute.

6. Corollary. If n is an integer, $(e^{\lambda})^n = e^{n\lambda}$ Moreover, $(e^{\lambda})^{-1} = e^{-\lambda}$.

7. Corollary. $e^{\lambda}\lambda = \lambda e^{\lambda}$.

Proof. We have

$$e^{\lambda}e^{t\lambda} = e^{t\lambda}e^{\lambda}\ (= e^{(t+1)\lambda}).$$

Hence, using Lemma 3, and taking derivatives,

$$e^{\lambda}e^{t\lambda}\lambda = e^{t\lambda}\lambda e^{\lambda}$$

METHODIST COLLEGE LIBRARY
Fayetteville, N. C.
47144

Now set $t = 0$. Q.E.D.

8. Corollary. The equation $\dfrac{dh}{dt} = \lambda h$ $(h(0) = e)$ has the unique solution $g = e^{t\lambda}$.

Corollary 8 follows immediately from Lemma 3 and Corollary 7.

The following lemma connects the exponential function, a homomorphism f, and the induced map f_* of the Lie algebras.

9. Lemma. Let $f: G \to G'$ be a homomorphism, and let f_* be the induced homomorphism of the Lie algebras $\Lambda \to \Lambda'$. Then, for any $\lambda \in \Lambda$,

$$f(e^{\lambda}) = e^{f_*(\lambda)} \tag{3}$$

Proof. Let $g(t) = e^{t\lambda}$. Then

$$\frac{dg}{dt} = g\lambda; \;\; g(0) = e.$$

We now apply f_* to both sides of the first equation and use Lemmas 1.21 and 3.5 to obtain

$$\frac{d}{dt}(fg(t)) = f_*\left(\frac{dg}{dt}\right) = f_*(g\lambda) = f(g(t))\,(f_*\lambda); \;\; fg(e) = e'.$$

Thus, fg satisfies the differential equation defining the exponential, and has the correct initial condition. Hence

$$f(g(t)) = e^{tf_*(\lambda)}$$

Now put $t = 1$. Q.E.D.

Remark. If f is a local homomorphism, equation (3) is valid for λ in some neighborhood of 0 in Λ. This is clear from the above proof and the remark following Lemma 3.5.

In the next section, we shall prove a "converse" of Lemma 9, showing that a homomorphism of Lie groups may be deduced from a homomorphism of their Lie algebras. Formula (3) gives the recipe which will be used.

If we deal with the Lie algebra of matrices, the exponential takes on its usual form. This is because (cf. Lemma 3.15) equation (1) reduces to the matrix equation

$$\frac{dS\,(t)}{dt} = S(t)\,\lambda; \;\; S(0) = I$$

It is very easily verified that the solution of this equation is

$$S(t) = I + t\lambda + \frac{1}{2!}t^2\lambda^2 + \cdots = \exp(t\lambda).$$

Hence, for matrices,

$$\exp \lambda = I + \lambda + \frac{1}{2!}\lambda^2 + \cdots \tag{4}$$

In this particular case, the various properties of the exponential given in Theorem 4 through Corollary 8 may be verified by a direct computation.

The following lemma records equation (4) for later reference.

10. Lemma. Let $f(z) = e^z$ for complex values of z, and let λ be any $n \times n$ matrix interpreted as an element of the Lie algebra of $GL(n)$. Then $\exp \lambda = f(\lambda) = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!}$

We now show how the exponential function gives a useful coordinate system for a Lie group.

11. Lemma. The mapping $\lambda \to \exp \lambda$ is a C^∞ homeomorphism of some neighborhood of 0 in Λ onto some neighborhood U of e in G.

Proof. We already know that the map $g : \lambda \to \exp \lambda$ is C^∞. We have $\left.\frac{d}{dt} e^{t\lambda}\right|_{t=0} = \lambda$, and thus by Definition 1.19 and Lemma 1.21, $f_*(\lambda) = \lambda$. Our assertion now follows from Lemma 1.18(c).

12. Definition. If U is a neighborhood of e as in Lemma 11, and $g \varepsilon U$, we define $\lambda = \log g$ as the inverse function of $g = e^\lambda$.

Remark. For matrices a neighborhood of I consists of all matrices $I + X$, where $|X| < 1$. These have inverses given by $(I + X)^{-1} = I - X + X^2 \ldots$. We may easily show that, $\log(I + X) = X - \frac{X^2}{2} + \frac{X^3}{3} - \cdots$. To see this, we note that $\log(1 + z) = z - \frac{z^2}{2} + \cdots$, and $\exp z = 1 + z + \frac{z^2}{2!} + \cdots$ are related by the formulas

$$\exp\log(1 + z) = 1 + z; \; \log\exp(z) = z \quad \text{for} \quad |z| < 1.$$

By Lemma 0.35, those formulas remain true for matrices. This shows that the log function defined by $\log(1 + X) = X - \frac{X^2}{2} + \frac{X^3}{3} \cdots$ is the inverse of the exponential function for $|X| < 1$.

Lemma 11 shows that the function $\log g$ is a *coordinate function* for a neighborhood of e. The coordinate space is the Lie algebra Λ. We formulate this as a definition.

13. Definition. Let G be a Lie group, Λ its Lie algebra. The *logarithmic coordinate* system of G is the coordinate system (p, U) about e, where $p: U \to \Lambda$, with $p(g) = \log g$ for $g \in U$ it being understood that U is chosen so that $\log g$ is single valued and C^∞ in U.

Remark. In this coordinate system, e^λ has the coordinate expression λ if λ is small enough. Further, the curves $g(t) = e^{t\lambda}$ are expressed by the lines $t\lambda$. Moreover, if $\lambda \varepsilon \Lambda$, the coordinate of λ is $\tilde{\lambda} = \lambda$. For, $e^{t\lambda}$ represents λ at $t = 0$. Hence $t\lambda$ is the coordinate expression for a curve representing λ, and differentiating, we obtain $\tilde{\lambda} = \lambda$.

Out of an infinity of possible coordinate systems, we distinguished one. We now show that our logarithmic coordinates are particularly useful. Until now we have always assumed homomorphisms to be C^∞. We now show that *continuity* of a homomorphism f is sufficient to ensure that f is C^∞.

14. Theorem. Let $f: G \to G'$ be a group homomorphism of the Lie groups G and G'. Let f be continuous. Then f is C^∞.

Proof. By Lemma 2.6, it is sufficient to prove that f is C^∞ at e. Let V and V' be spherical neighborhoods of 0 in Λ and Λ' such that if we write $2V = \{\lambda | \lambda = 2\lambda_1, \lambda_1 \in V\}$, etc., then the exponential is a homeomorphism from $2V$ onto a neighborhood of e in G, and similarly a homeomorphism from $2V'$ onto a neighborhood of e' in G'. If $U = \exp V$, $U' = \exp V'$, we assume V is so small that $f(U) \subset U'$. It then follows that for $\lambda \in V$, there is a unique $\mu \in V'$ such that $f(\exp \lambda) = \exp \mu$. We maintain that $f(\exp(\lambda/2)) = \exp(\mu/2)$. For, suppose $f(\exp(\lambda/2)) = \exp \mu'$. Then $\exp 2\mu' = [f(\exp(\lambda/2))]^2 = f[\exp(\lambda/2)^2] = f(\exp \lambda) = \exp \mu$. Thus, $2\mu' = \mu$, which is what was asserted. It follows by induction that $f(\exp(2^{-n}\lambda)) = \exp(2^{-n}\mu)$. Therefore, using the fact that f is a homomorphism and the known properties of the exponential, we have

$$f(\exp(k\, 2^{-n}\lambda)) = [f(\exp(2^{-n}\lambda))]^k = \exp[k \cdot 2^{-n}\mu]$$

for any integer k. That is, $f(\exp(t\lambda)) = \exp(t\mu)$ for each dyadic rational t. By continuity, this same formula must hold for each real t. We have therefore shown that for each $\lambda \in V$ there is a $\mu \in V'$ such that

$$f(\exp t\lambda) = \exp(t\mu)$$

Next let $\lambda_1, \ldots, \lambda_n$ be a basis for Λ, $\lambda_i \in V$, and let $\mu_1, \ldots, \mu_n$ satisfy $f(\exp t\lambda_i) = \exp(t\mu_i)$. We now define a mapping of n-space into G by the formula

$$\eta(t^1, \ldots, t^n) = \exp(t^1\lambda_1) \cdots \exp(t^n\mu_n).$$

The mapping η is clearly C^∞. In logarithmic coordinates η is expressed near $t = 0$ by the formula

$$\eta(t) \sim t^1\lambda_1 + \cdots + t^n\lambda_n + \cdots$$

since the exponential $\exp(t^i\lambda_i)$ is expressed by $t^i\lambda_i$ and multiplication is always given by the sum, to first order. Since $\lambda_1, \ldots, \lambda_n$ are independent, this shows that the linear part of η is non-singular and therefore by the implicit function theorem (Lemma 1.18c) η maps some neighborhood V^1 of R^n onto a neighborhood U^1 of e in G, and η and η^{-1} are C^∞. Now

$$\begin{aligned} f(\eta(t)) &= f(\exp(t^1\lambda_1) \cdots \exp(t^n\lambda_n)) \\ &= f(\exp(t^1\lambda_1) \cdots f(\exp(t^n\lambda_n)) \\ &= \exp(t^1\mu_1) \cdots \exp(t^n\mu_n) = \eta_1(t) \end{aligned}$$

where η_1 is also C^∞. But this formula shows $f\eta = \eta_1$. Hence $f = \eta_1\eta^{-1}$ in the neighborhood U^1 of e in G, and f is C^∞ in U^1. Q.E.D.

Remark. It is actually sufficient for Theorem 14 to assume that f is continuous only at e.

5. The Campbell-Baker-Hausdorff Theorem

The theorem whose name is the title of this section gives the formula for multiplication in a Lie group in terms of the logarithmic coordinates on the Lie group. Recall that if $\lambda, \mu \in \Lambda$ and λ, μ are sufficiently close to 0, then λ and μ are the logarithmic coordinates of $g = e^\lambda$ and $h = e^\mu$. In this coordinate system the multiplication function $m(\lambda, \mu)$ is given by

$$e^\lambda e^\mu = e^{m(\lambda,\mu)}; \; m(\lambda, \mu) = \log(e^\lambda e^\mu)$$

From general considerations, we know that $m(\lambda, \mu) = \lambda + \mu + \cdots$. The Campbell-Baker-Hausdorff formula gives a complete analytic expression for $m(\lambda, \mu)$ in terms of iterated Lie brackets of the elements λ and μ. For example, the quadratic term is $\frac{1}{2}[\lambda, \mu]$.

To avoid circumlocution in stating the formula of Campbell, Baker, and Hausdorff, we introduce some useful notations.

1. Definition. Let Λ be a Lie algebra, $\lambda \in \Lambda$. Then Ad λ is the linear transformation in Λ defined by the formula

$$(\mathrm{Ad}\,\lambda)(\mu) = [\lambda, \mu]$$

Note, e.g., that $(\mathrm{Ad}\,\lambda)^3 \mu = [\lambda, [\lambda, [\lambda, \mu]]]$.

Warning: If Λ is the Lie algebra of all $n \times n$ matrices, Ad λ must be regarded as an $n^2 \times n^2$ matrix, since it operates on Λ – an n^2 dimensional space.

According to Lemma 3.15, the set of all linear transformations on the space Λ form a Lie algebra with the bracket operation $[S, T] = ST - TS$. The reader will easily verify that the Jacobi identity implies the following statement.

2. Lemma. The mapping $\lambda \to \mathrm{Ad}\,\lambda$ is a Lie algebra homomorphism.

3. Lemma. Let Λ be the Lie algebra of a Lie group G, and let $\lambda, \mu \in \Lambda$. Then

$$e^{\lambda}\mu e^{-\lambda} = (\exp(\mathrm{Ad}\,\lambda))\,\mu \left(= \mu + [\lambda, \mu] + \frac{1}{2!}[\lambda, [\lambda, \mu]] + \cdots \right)$$

Proof. Let

$$\mu(t) = e^{t\lambda}\mu e^{-t\lambda}; \quad \mu(0) = \mu \tag{1}$$

For each t_0 and t, we clearly have

$$\mu(t_0 + t) = e^{(t_0+t)\lambda}\mu e^{-(t_0+t)\lambda} = e^{t\lambda}\mu(t_0)\,e^{-t\lambda}$$

Therefore, by Definition 3.7, we have

$$\dot{\mu}(t_0) = \frac{d}{dt}\mu(t_0 + t)\bigg|_{t=0} = \frac{d}{dt}(e^{t\lambda}\mu(t_0)\,e^{-t\lambda})\bigg|_{t=0} = \left[\frac{d}{dt}e^{t\lambda}\bigg|_{t=0}, \mu(t_0)\right]$$

$$= [\lambda, \mu(t_0)] = (\mathrm{Ad}\,\lambda)(\mu(t_0))$$

Thus, $\mu(t)$ satisfie the differential equation

$$\frac{d\mu(t)}{dt} = (\mathrm{Ad}\ \lambda)\ \mu(t); \mu(0) = \mu \tag{2}$$

But this equation is readily seen to have the solution

$$\begin{aligned} \mu(t) &= \mu + t(\mathrm{Ad}\ \lambda)\ \mu + \frac{t^2(\mathrm{Ad}\ \lambda)^2}{2!}\ \mu + \cdots \\ &= (\exp (t\ \mathrm{Ad}\ \lambda))\ (\mu) \end{aligned} \tag{3}$$

Hence, finally, using formulas (1) and (3), we have

$$e^{\lambda}\mu e^{-\lambda} = \mu(1) = (\exp (\mathrm{Ad}\ \lambda))\ \mu$$

This is the required result. Q.E.D.

The next lemma shows how e^{λ} evolves as λ varies over an arbitrary path.

4. Lemma. Let $\lambda(t)$ by a C^{∞} curve in Λ, the Lie algebra of G. Let

$$\varphi(z) = (e^z - 1)/z = 1 + \frac{1}{2!}z + \frac{1}{3!}z^2 + \cdots \tag{4}$$

Then

$$e^{-\lambda(t)}\frac{d}{dt}e^{\lambda(t)} = \varphi(-\mathrm{Ad}\ \lambda(t))\ (\lambda'(t)) \tag{5}$$

Proof. Let $g(s, t) = e^{s\lambda(t)}$, and let

$$\alpha(s, t) = g^{-1}\frac{\partial g}{\partial s} = e^{-s\lambda(t)}e^{s\lambda(t)}\lambda(t) = \lambda(t)$$

$$\beta(s, t) = g^{-1}\frac{\partial g}{\partial t} = e^{-s\lambda(t)}\frac{\partial}{\partial t}e^{s\lambda(t)} \tag{6}$$

Then, by Lemma 3.14, we have

$$\frac{\partial \beta}{\partial s} - \frac{\partial \alpha}{\partial t} + [\alpha, \beta] = 0.$$

Hence

$$\frac{\partial \beta}{\partial s} - \lambda'(t) + [\lambda(t), \beta] = 0.$$

We now fix t, set $\lambda = \lambda(t)$, $\lambda' = \lambda'(t)$, and write the last equation as follows:

$$\frac{d\beta}{ds} = (-\mathrm{Ad}\,\lambda)\,\beta + \lambda'; \;\; \beta(0) = 0 \tag{7}$$

where the initial condition follows from (6). But equation (7) may be solved explicitly in terms of an everywhere convergent power series in s:

$$\beta(s, t) = \beta(s) = s\lambda' + \frac{s^2}{2}(-\mathrm{Ad}\,\lambda)\,\lambda' + \cdots + \frac{s^n}{n!}(-\mathrm{Ad}\,\lambda)^{n-1}\,\lambda' + \cdots$$

Putting $s = 1$, and using formula (4), we have

$$\beta(1, t) = \beta(1) = \varphi(-\mathrm{Ad}\,\lambda)\,\lambda'$$

Using formula (6), we have the stated result. Q.E.D.

5. Theorem. (Campbell-Baker-Hausdorff.) Let Λ be the Lie algebra of a Lie group G. Then for λ, μ in a sufficiently small neighborhood of 0 in Λ, we have

$$\log(e^{\lambda}e^{\mu}) = \lambda + \int_0^1 \psi[(\exp \mathrm{Ad}\,(\lambda))\,(\exp t\,\mathrm{Ad}\,(\mu)]\,(\mu)\,dt$$

where $\psi(z) = (z \log z)/(z - 1)$ is analytic near $z = 1$.

Proof. Let U be a neighborhood of e in which the logarithm exists, and let V be a neighborhood of e so small that for $g, h \in V$ we have $gh \in U$. Finally, let N be a spherical neighborhood of 0 in Λ such that if $\lambda \in N$ then $e^{\lambda} \in V$. We now fix $\lambda, \mu \in N$, and put

$$\Gamma = \Gamma(t) = \Gamma(t, \lambda, \mu) = \log(e^{\lambda}e^{t\mu}).$$

Then, by the definition of the logarithm,

$$e^{\lambda}e^{t\mu} = e^{\Gamma(t)} \tag{8}$$

We now compute $\mathrm{Ad}\,\Gamma(t)$ in terms of $\mathrm{Ad}\,\lambda$ and $\mathrm{Ad}\,\mu$. For each $\nu \in \Lambda$, we have

$$e^{\Gamma(t)}\,\nu\,e^{-\Gamma(t)} = e^{\lambda}(e^{t\mu}\,\nu\,e^{-t\mu})\,e^{-\lambda} \tag{9}$$

Thus, by Lemma 3, we have

$$(\exp \mathrm{Ad}\,\Gamma)\,\nu = (\exp \mathrm{Ad}\,\lambda)\,(\exp \mathrm{Ad}\,t\mu)\,\nu$$

and therefore

$$\exp \operatorname{Ad} \Gamma = (\exp \operatorname{Ad} \lambda)(\exp t \operatorname{Ad} \mu) \tag{10}$$

We now assume that λ and μ are sufficiently small so that $|(\exp \operatorname{Ad} \lambda)(\exp t \operatorname{Ad} \mu) - I| < 1$ for $0 \leqslant t \leqslant 1$. Then we may take logarithms in (10) and we have

$$\operatorname{Ad} \Gamma = \log [(\exp \operatorname{Ad} \lambda)(\exp t \operatorname{Ad} \mu)] \tag{11}$$

Differentiating (8), we have

$$\frac{d}{dt} e^{\Gamma(t)} = \frac{d}{dt}(e^{\lambda} e^{t\mu}) = e^{\lambda} e^{t\mu} \mu$$

$$\frac{d}{dt} e^{\Gamma(t)} = e^{\Gamma(t)} \mu$$

$$e^{-\Gamma(t)} \frac{d}{dt} e^{\Gamma(t)} = \mu$$

Hence, by Lemma 4, we have

$$\varphi(-\operatorname{Ad} \Gamma(t))\, \Gamma'(t) = \mu \tag{12}$$

where $\varphi(z) = (e^z - 1)/z$. If $|z - 1| < 1$, we have $\varphi(-\log z) = (z - 1)/(z \log z)$. Thus, setting $\psi(z) = (z \log z)/(z - 1)$, we have

$$\psi(z)\, \varphi(-\log z) = 1 \tag{13}$$

We now substitute the linear transformation $(\exp \operatorname{Ad} \lambda)(\exp t \operatorname{Ad} \mu)$ for z in (13) and use formula (11):

$$\psi[(\exp(\operatorname{Ad} \lambda))(\exp(t \operatorname{Ad} \mu))]\, \varphi(-\operatorname{Ad} \Gamma(t)) = \mathrm{I} \tag{14}$$

This equation gives the inverse of $\varphi(-\operatorname{Ad} \Gamma(t))$ and permits rewriting (12) in the form

$$\Gamma'(t) = \psi[(\exp \operatorname{Ad} \lambda)(\exp t \operatorname{Ad} \mu)](\mu) \tag{15}$$

Finally, we integrate (15) from 0 to 1 and recall that $\Gamma(1) = \log(e^{\lambda} e^{\mu})$ and $\Gamma(0) = \lambda$, thereby obtaining the result. Q.E.D.

Remark. The first few terms are

$$\log(e^{\lambda} e^{\mu}) = \lambda + \mu + \tfrac{1}{2}[\lambda, \mu] + \tfrac{1}{12}[\lambda, [\lambda, \mu]] - \tfrac{1}{12}[\mu, [\mu, \lambda]] + \cdots$$

The Campbell-Baker-Hausdorff formula gives a direct relationship between the Lie algebra of a group and the local structure of that group. For the remainder of the present section, we exploit this relationship. We first give a converse of Lemma 3.12.

6. Theorem. Let Λ and Λ' be the Lie algebras of the Lie groups G and G' and let $\sigma : \Lambda \to \Lambda'$ be a Lie algebra homomorphism. Then there exists a local Lie group homomorphism f sending some neighborhood of e in G into G', and such that $f_* = \sigma$. Any two such local homomorphisms agree in some neighborhood of e.

Proof. According to Lemma 4.9, if such a local homomorphism f existed, we would have

$$f(e^{\lambda}) = e^{\sigma(\lambda)} \tag{1}$$

for λ in some neighborhood of 0 in Λ. Since the mapping $\lambda \to e^{\lambda}$ is a homeomorphism of a neighborhood of 0 in Λ onto a neighborhood of e in G, it follows that f is uniquely defined in some neighborhood of e in G. To show that f exists, we let U be a neighborhood of e in G in which the logarithm exists, and put

$$f(g) = \exp\,(\sigma(\log g)) \quad g \in U. \tag{2}$$

This makes (1) valid for small λ.

We now verify that $f_* = \sigma$ and that f is a local homomorphism. We go to logarithmic coordinates in G and G', and suppose that $m(\lambda, \mu)$ and $m'(\lambda', \mu')$ express multiplication in G and G' respectively. By (1), we have, simply,

$$\tilde{f}(\lambda) = \sigma(\lambda) \tag{3}$$

for every λ. Thus, σ is the "linear part" of $\tilde{f}$. By Lemma 1.16, $\widetilde{f_*(\lambda)} = \sigma(\tilde{\lambda})$. But in logarithmic coordinates, $\tilde{\lambda} = \lambda$, and therefore $f_*(\lambda) = \sigma(\lambda)$, which verifies $f_* = \sigma$.

To prove that f is a local homomorphism, we must prove that $\tilde{f}(m(\lambda, \mu)) = m'(\tilde{f}(\lambda), \tilde{f}(\mu))$, i.e. $\sigma(m(\lambda, \mu)) = m'(\sigma\lambda, \sigma\mu)$. But according to the Campbell-Baker-Hausdorff formula,

$$m(\lambda, \mu) = \int_0^1 \psi[(\exp \operatorname{Ad} \lambda)\,(\exp \operatorname{Ad} t\mu)]\,(\mu)\, dt + \lambda$$

where $\psi(z)$ is a certain function of z analytic near $z = 1$. Now, since σ preserves Lie brackets, we have $\sigma[(\operatorname{Ad} \lambda_1)\,(\lambda_2)] = [\operatorname{Ad}\,(\sigma\lambda_1)]\,(\sigma\lambda_2)$. By repeated application of this result and use of power series (recalling that σ is linear and hence also continuous), we have

$$\begin{aligned}\sigma[m(\lambda, \mu)] &= \sigma \int_0^1 \psi[(\exp \operatorname{Ad} \lambda)(\exp \operatorname{Ad} t\mu)](\mu)\, dt + \lambda \\ &= \int_0^1 \psi[(\exp \operatorname{Ad}(\sigma\lambda))(\exp \operatorname{Ad} t(\sigma\mu)](\sigma\mu)\, dt + \sigma\lambda \\ &= m'(\sigma\lambda, \sigma\mu)\end{aligned}$$

for all sufficiently small λ and μ, and our proof is complete. Q.E.D.

7. Corollary. Let Λ and Λ' be the Lie algebras of G and G' respectively. Then Λ is isomorphic to Λ' if and only if G and G' are locally (C^∞) isomorphic. If G and G' are simply connected then Λ and Λ' are isomorphic if and only if G and G' are C^∞ isomorphic.

The proof is an easy consequence of the preceding theorem, Lemma 3.12, Theorem 2.7 and Corollary 2.8. The following corollary is also immediate from Corollary 2.8.

8. Corollary. If Λ, Λ' are the Lie algebras of the Lie groups G, G' and if G is simply connected, then the correspondence $f \to f_* = \sigma$ is a $1-1$ onto correspondence between C^∞ group homomorphism $f: G \to G'$ and Lie algebra homomorphisms $\sigma: \Lambda \to \Lambda'$.

In section II.7.12, we shall prove the "global existence theorem" that every (abstract) Lie algebra is (isomorphic to) the Lie algebra Λ of a Lie group G. By Corollary 2.7, we see that this puts the family of simply connected Lie groups in a natural $1-1$ onto correspondence with the family of abstract Lie algebras.

Remark. The Campbell-Baker-Hausdorff formula shows that when logarithmic coordinates are used in the neighborhood of e, the multiplication function $m(\lambda, \mu)$ is *analytic*—not merely C^∞. If (p, U) is a logarithmic coordinate system, we may cover all of G with the translates (pg^{-1}, gU) of this coordinate system. It is fairly clear, and simple to prove, that if these coordinate systems are used on G, then G becomes an *analytic manifold*, that is, the change of coordinates in overlapping coordinate systems is analytic. Further, multiplication and the taking of inverses become *analytic*. It is then possible to recast all of our results and strengthen C^∞ conclusions to analytic ones. However, we shall have no need for these results.

6. Subgroups and Sub-algebras

In this section, we relate properties of Lie subgroups of a Lie group to properties of corresponding Lie sub-algebras of its Lie algebra. As before, we let G denote a Lie group and $\Lambda = \Lambda(G)$ its Lie algebra.

1. Lemma. Let H be a Lie subgroup of G (see definition 2.12) with Lie algebra $\Lambda(H)$. Then the bracket operation in $\Lambda(H)$ is induced from the bracket product in $\Lambda = \Lambda(G)$ by the natural embedding of $\Lambda(H)$ in $\Lambda(G)$ given in Lemma 1.30. Hence, $\Lambda(H)$ is closed under the bracket operation; i.e., if $\lambda, \mu \in \Lambda(H) \subseteq \Lambda$, then $[\lambda, \mu] \in \Lambda(H)$.

Proof. Let $i : H \to G$ be the inclusion map defined by $i(h) = h$. The embedding as described in Lemma 1.30 identifies $\lambda \in \Lambda(H)$ with $i_*\lambda \in \Lambda(G) = \Lambda$. Therefore the result we wish to prove is $[i_*\lambda, i_*\mu] = i_*[\lambda, \mu]$ for λ, $\mu \in \Lambda(H)$. But i is clearly a homomorphism of H into G, and hence the present lemma follows from Lemma 3.12. Q.E.D.

2. Definition. Let Λ be a Lie algebra, and let $\Lambda' \subseteq \Lambda$. Λ' is called a *Lie sub-algebra*, or simply a *sub-algebra* of Λ if it is a Lie algebra under the operations in Λ; i.e., if Λ' is a vector subspace of Λ which is also closed under the Lie bracket operation of Λ.

Lemma 1 may therefore be restated, "If H is a Lie sub-group of G, then $\Lambda(H)$ is a Lie sub-algebra of $\Lambda(G)$." We now show that the Lie sub-algebras of $\Lambda(G)$ determine all the connected Lie subgroups of G.

3. Theorem. Let G be a Lie group with Lie algebra Λ. Then the correspondence $H \to \Lambda(H)$ is a $1 - 1$ onto correspondence between the class of all the connected Lie subgroups H of G and the class of all sub-algebras of Λ.

Proof. Lemma 1 shows that if H is a Lie subgroup of G then $\Lambda(H)$ is a sub-algebra of $\Lambda = \Lambda(G)$. We first show that this correspondence is onto.

Let Λ' be a sub-algebra of Λ. Using the Campbell-Baker-Hausdorff theorem, we shall construct a local subgroup L of G (cf. Definition 2.13). We take Λ' as the coordinate space for L and we let $g: \Lambda' \to G$ according to the formula

$$g(\lambda') = \exp \lambda'$$

We know already that g is regular in some neighborhood of 0 in Λ and hence in some spherical neighborhood V' of 0 in Λ'. Thus, g injects the manifold V' into G, and $g(V')$ is a submanifold L of G (cf. Definition 1.28). We now show that L is a local subgroup of G. To do this, we express multiplication in the coordinate space V' using the Campbell-Baker-Hausdorff theorem:

$$m(\lambda', \mu') = \lambda' + \mu' + \tfrac{1}{2}[\lambda', \mu'] + \cdots$$

where "..." involves iterated bracket operations on λ' and μ' and the series converges for λ', μ' in some neighborhood V'' of 0 in V'. Since Λ' is closed under the bracket operation and is also closed under the operation of taking limits (it is a finite dimensional vector space), we see that if V'' is small enough, the C^∞ function $m(\lambda', \mu')$ maps $V'' \times V''$ into V'. This shows that L is a local subgroup of G with germ $U = g(V'')$. By Theorem 2.14, V is a neighborhood of e in some unique connected submanifold H. The Lie algebra of H—which is to say, the tangent space at e—is clearly Λ'.

To show that the correspondence $H \to \Lambda(H)$ is $1-1$, suppose that the connected Lie subgroup H' had the same Lie algebra Λ' as the subgroup H which was constructed above. Then for some $\delta > 0$, the mapping $\lambda' \to \exp \lambda' = g(\lambda')$ would be a C^∞ homeomorphism for all $\lambda' \in \Lambda'$ satisfying $|\lambda'| < \delta$. Thus, H' would contain a neighborhood of e which was also a neighborhood of e in H, and therefore $H = H'$ by uniqueness result stated in Theorem 2.14.

(Note that we have used the obvious fact that if $\lambda \in \Lambda(H) \subseteq \Lambda(G)$, then $\exp \lambda \in H \subseteq G$ is uniquely determined by the general definition of the exponential function regardless of whether λ is considered to be an element of $\Lambda(H)$ or of $\Lambda(G)$. Indeed, if $\lambda \in \Lambda(H)$, we have $\exp(i_*\lambda) = i(\exp \lambda)$ by Lemma 4.9.) Q.E.D.

We now show how certain properties of subgroups (and groups) may be determined from corresponding properties of their sub-algebras (and algebras).

4. Definition. Let Λ' be a sub-algebra of the Lie algebra Λ. Suppose that $[\Lambda, \Lambda'] \subseteq \Lambda'$; i.e. if $\lambda \in \Lambda$, $\lambda' \in \Lambda'$, then $[\lambda, \lambda'] \in \Lambda'$. Then Λ' is called an ideal of Λ.

5. Definition. Let H be a Lie subgroup of G. H is an *invariant* subgroup, or *normal* subgroup, if for $h \in H$, $g \in G$ we have $ghg^{-1} \in H$ and if the map $(g, h) \to ghg^{-1}$ is C^∞ from $G \times H$ onto H.

It follows from this definition that if H is an invariant subgroup of G, and if $\lambda' \in \Lambda(H)$ while $g \in G$, we have $g\lambda' g^{-1} \in \Lambda(H)$ and the map $g \to g\lambda' g^{-1}$ is a C^∞ mapping from G into $\Lambda(H)$. For if $\lambda' \in \Lambda(H)$ and if $\lambda' = \left.\frac{dh}{dt}\right|_{t=0}$, we have $g\lambda' g^{-1} = \left.\frac{d}{dt} ghg^{-1}\right|_{t=0} \in \Lambda(H)$ since $gh(t)\, g^{-1}$ is a curve in H. We now relate Definitions 4 and 5.

6. Lemma. Let H be a connected Lie subgroup of the connected Lie group G. Then H is an invariant subgroup of G if and only if $\Lambda(H)$ is an ideal of $\Lambda = \Lambda(G)$.

Proof. Suppose, first, that H is an invariant subgroup of G. If $\mu \in \Lambda(H)$ and $g \in G$, it follows that $g\mu g^{-1} \in \Lambda(H)$. Now if $\lambda \in \Lambda$, then $e^{\lambda t}\mu e^{-\lambda t}$ is a C^∞ curve in $\Lambda(H)$ and therefore $[\lambda, \mu] = \left.\frac{d}{dt}(e^{\lambda t}\mu e^{-\lambda t})\right|_{t=0} \in \Lambda(H)$.

Conversely, suppose $\Lambda' = \Lambda(H)$ is an ideal in $\Lambda = \Lambda(G)$. By the use of the Campbell-Baker-Hausdorff theorem (Theorem 5.5) and equation (8) of Lemma 3.7, we may compute

$$\log (e^{\lambda} e^{\lambda'} e^{-\lambda}) = \lambda' + [\lambda, \lambda'] + \cdots$$

for $|\lambda|$, $|\lambda'|$ sufficiently small. If $\lambda \in \Lambda$, $\lambda' \in \Lambda'$, this formula shows that $\log (e^{\lambda} e^{\lambda'} e^{-\lambda}) \in \Lambda'$ and therefore $e^{\lambda} e^{\lambda'} e^{-\lambda} \in H$. Since $\lambda \to \exp \lambda$ ($\lambda' \to \exp \lambda'$) is a homeomorphism near $\lambda = 0$, there exists a neighborhood U of e in G and a neighborhood U' of e in H such that $gh'g^{-1} \in H$ for $h' \in U'$ and $g \in U$. Since $gh_1'h_2' \cdots h_n'g^{-1} = (gh_1'g^{-1}) \cdots (gh_n'g^{-1})$, and since U' generates H, it follows that $ghg^{-1} \in H$ for $h \in H$, $g \in U$. Finally, since U generates G, and since $(g_1 \cdots g_n)\, h(g_1 \cdots g_n)^{-1} = g_1(g_2 \cdots g_n h(g_2 \cdots g_n)^{-1}) g_1^{-1}$ we see by induction on n that $ghg^{-1} \in H$ for $g \in G$, $h \in H$; i.e. that H is an invariant subgroup in the sense of abstract group theory. It is plain from the above argument that the mapping $(g, h) \to ghg^{-1}$ is C^∞ from $G \times H$ into H, and this completes the proof. Q.E.D.

The correspondence given in Lemma 6 between ideals and invariant Lie subgroups suggests a correspondence between quotient groups and quotient algebras.

7. Definition. Let Λ be a Lie algebra, and let I be an ideal in Λ. The *factor algebra* or *quotient algebra* Λ/I is defined as the set of all cosets $\lambda + I$ made into a Lie algebra by the usual algebraic method:

$$(\lambda + I) + (\mu + I) = (\lambda + \mu) + I$$

$$c(\lambda + I) = c\lambda + I$$

$$[\lambda + I, \mu + I] = [\lambda, \mu] + I$$

Remark. Since I is an ideal it follows easily that these definitions are single-valued and that Λ/I is a Lie algebra. The standard 'homomorphism theorems' of abstract are also clearly valid in the context of Lie algebra theory: If $\sigma: \Lambda \to \Lambda'$ is a Lie algebra homomorphism with kernel $I = \sigma^{-1}(0)$, then I is an ideal of Λ, $\sigma(\Lambda)$ is a sub-algebra of Λ' and there is a natural isomorphism between the image Lie algebra $\sigma(\Lambda)$ and Λ/I.

8. Lemma. Let G be a Lie group and let N be a closed normal Lie subgroup (cf. Definition 5). Let $G/N = G'$ be the factor group of G by the subgroup N and let f be the natural map of G into G', i.e. the map which sends each $g \in G$ into its coset gN. Then we may define a topology for G', a family of open sets U_α covering G' and a family of coordinate systems (P_α, U_α) in such a way that G' becomes a Lie group, that f is a C^∞ map of G onto G' which maps open sets onto open sets, and that f_* maps the Lie algebra of G onto the Lie algebra of G'.

This statement should be compared to Theorem 2.11. We omit the proof which is quite similar to the proof of that theorem, except for the details concerning the C^∞ structure of G'. We note that it is essential that N be closed. For $N = f^{-1}(e)$ must be closed if f is to be continuous since the inverse image of a point under a continuous mapping is a closed set.

Recall the example immediately preceding Definition 1.28 of a normal one-dimensional subgroup which was dense in an n-dimensional torus—hence not closed. We now show that such pathological normal subgroups do not occur when G is simply connected, so that for simply connected groups, the correspondence between ideals and normal subgroup, and between quotient algebras and quotient groups is complete.

9. Theorem. Let G be a simply connected group, and let N be a connected invariant Lie subgroup. Then N is closed. If I is the Lie algebra of N, then the Lie group G/N has as Lie algebra Λ/I. Thus, there is a $1 - 1$ correspondence between quotient algebras Λ/I and quotient groups G/N for which N is a connected normal Lie subgroup of G.

Proof. Let Λ be the Lie algebra of G, $I \subseteq \Lambda$ the Lie algebra of N. Then I is an ideal by Lemma 6 and we let $\Lambda' = \Lambda/I$ be the quotient Lie algebra. We now use the theorem—which we will prove much later in the present book—that every Lie algebra is the Lie algebra of some Lie group (cf. Theorem II.7.12). Let G' be a group whose Lie algebra is (isomorphic to) Λ'. Now let $\sigma: \Lambda \to \Lambda/I$ be the natural Lie algebra homorphism defined by $\sigma(\lambda) = \lambda + I$. According to Theorem 5.6 there is a local homomorphism f from G into G' with $f_* = \sigma$. Using Theorem 2.7, we may extend f to be a homomorphism in the large from G into G'. We put $\hat{N} = f^{-1}(e)$, and we first show that $N \subseteq \hat{N}$. To see this, note that for λ sufficiently small with $\lambda \in I$ we have $f(e^\lambda) = e^{f_*\lambda} = e^{\sigma\lambda} = e^0 = e$ (cf. Lemma 4.9). Thus a neighborhood V of e in N satisfies $V \subseteq \hat{N}$. Since N is connected, V generates N, and since $\hat{N}$ is a subgroup, we have $N \subseteq \hat{N}$.

We can now show that N is closed. Take U' to be a neighborhood of the identity in G' in which the logarithm exists. Let L be a spherical neighborhood of 0 in Λ so small that $\lambda \to \exp \lambda$ is a homeomorphism from L onto a neighborhood U of e in G, such that $f(U) \subseteq U'$, so that $\exp \sigma L \subseteq U'$. We maintain that $\hat{N} \cap U \subseteq N$. For if $g \in \hat{N} \cap U$, we have $g = e^\lambda$, $\lambda \in L$, with $e = f(g) = e^{f_*\lambda} = e^{\sigma\lambda}$. Hence $\sigma\lambda = 0$ since the logarithm is unique in U', and hence $\sigma\lambda = 0$ or $\lambda \in I$ (by definition of σ); consequently $g = e^\lambda \in N$ since I is the Lie algebra of N. Thus, $\hat{N} \cap U \subseteq N$. To show that N is closed in G, we first note that $\hat{N} = f^{-1}(e)$ is closed in G. Now let $g_0 \notin N$. If we first suppose $g_0 \in \hat{N}$, then we maintain $g_0 U$ does not meet N. For supposing the contrary we would have $g_0 u = n$ where $u \in U$, $n \in N$, and thus $g_0^{-1}n = u \in U$ and $g_0^{-1}n \in \hat{N}$ (since $g_0^{-1} \in \hat{N}$, $n \in N \subseteq \hat{N}$). Thus $g_0^{-1}n \in U \cap \hat{N} \subseteq N$. Set $n_1 = g_0^{-1}n \in N$. Then $g_0 = nn_1^{-1} \in N$ which is a contradiction. If, on the other hand, $g_0 \notin \hat{N}$, then there exists a neighborhood V of g_0 such that $V \cap \hat{N} = \phi$ ($\hat{N}$ is closed). Since $N \subseteq \hat{N}$ we have $V \cap N = \phi$. Hence in any case there is a neighborhood of $g_0 \notin N$ which does not meet N; and therefore N is closed.

We must now show that G/N has the Lie algebra $\Lambda' = \Lambda/I$. (I is the Lie algebra of N.) To see this, we let $\hat{f}$ be natural map $f: G \to G/N$ given in Lemma 8. We have $\hat{f}(e^\lambda) = e^{\hat{f}_*\lambda}$ for $\lambda \in \Lambda$. We now prove that $\hat{f}_*\lambda = 0$ if and only if $\lambda \in I$. Since $\hat{f}_*$ is onto this will show (by the algebraic isomorphism theorem cited in the remark following Definition 7), that $\Lambda(G/N)$ is isomorphic in a natural way with $\Lambda(G)/I$.

Suppose then $\lambda \in I$. Then $e^{t\lambda}$ is a curve in N and hence $\hat{f}(e^{t\lambda}) = e^0$, or $e^\mu = e^{t\hat{f}_*(\lambda)}$. Taking the derivative at $t = 0$, we have $\hat{f}_*\lambda = 0$. Conversely, if $\hat{f}_*\lambda = 0$, we have $\hat{f}(e^{t\lambda}) = e^{t\hat{f}_*(\lambda)} = e$. Hence $\exp t\lambda \in N$. Taking derivatives at $t = 0$, we obtain $\lambda \in I$. Theorem 9 is thereby proved. Q.E.D.

Another aspect of the correspondence between properties of groups and properties of their Lie algebra is seen in the following results.

10. Definition. A Lie algebra Λ is *commutative* or *abelian* if $[\lambda, \mu] = 0$ for all $\lambda, \mu \in \Lambda$.

11. Theorem. If G is a commutative Lie group, then its Lie Algebra Λ is commutative. Conversely, if the Lie algebra Λ of G is commutative, and if G is connected, then G is commutative.

Proof: First, suppose G is commutative. Let

$$\lambda = \left.\frac{dg}{dt}\right|_{t=0}, \quad \mu = \left.\frac{dh}{dt}\right|_{t=0}$$

be arbitrary elements of Λ. Then for fixed t we have,

$$g(t)\,\mu g^{-1}(t) = \frac{d}{ds}\,g(t)\,h(s)\,g^{-1}(t)\Big|_{t=0} = \left.\frac{dh}{ds}\right|_{s=0} = \mu$$

Hence

$$[\lambda,\mu] = \frac{d}{dt}\,(g(t)\,\mu g^{-1}(t))\Big|_{t=0} = \frac{d}{dt}\,\mu\Big|_{t=0} = 0$$

and Λ is commutative.

Conversely, if Λ is commutative, it follows from the Campbell-Baker-Hausdorff theorem that

$$e^{\lambda}e^{\mu} = e^{\lambda+\mu+{}^1\!/_2[\lambda,\mu]+\cdots} = e^{\lambda+\mu}$$

for λ, μ sufficiently small in Λ. Hence $e^{\lambda}e^{\mu} = e^{\mu}e^{\lambda}$ and therefore $gh = hg$ for h, g in a sufficiently small neighborhood of e. But by Lemma 2.2, a neighborhood of e generates G if G is connected. But if the generators of a group commute, the group is commutative, giving the required converse. Q.E.D.

The following theorem shows that in the special case of closed subgroups, the possible pathologies of subgroups do not occur.

12. Theorem. Let H be a subgroup of the Lie group G (in the algebraic sense), and suppose that H is closed in G. Then H may be given a C^∞ structure in such a way as to make H into a Lie subgroup of G, whose topology is the subspace topology. The C^∞ structure of H is uniquely determined by this latter condition.

Remark. In brief, a closed subgroup of a Lie group is a Lie subgroup.

Proof. We let Λ_1 be the set of $\lambda \in \Lambda$ such that $\exp t\lambda \in H$ for all t. Since H is a subgroup this condition is implied either by the statement $\exp t\lambda \in H$

for all t satisfying $|t| < \delta$ for some $\delta > 0$, or by the statement $\exp t\lambda \in H$ for all positive t. We first show that Λ_1 is a subalgebra of Λ. It is clear that if $\lambda \in \Lambda_1$ and if c is any scalar, then $c\lambda \in \Lambda_1$. To show that Λ_1 is closed under addition, we note that for small t

$$\log(\exp t\lambda \exp t\mu) = t(\lambda + \mu) + O(t^2)$$

Thus,

$$\log\left(\exp \frac{t}{n} \lambda \exp \frac{t}{n} \mu\right) = \frac{t}{n}(\lambda + \mu) + O\left(\frac{1}{n^2}\right) \text{ as } n \to \infty$$

and

$$\log\left(\exp \frac{t}{n} \lambda \exp \frac{t}{n} \mu\right)^n = t(\lambda + \mu) + O\left(\frac{1}{n}\right) \text{ as } n \to \infty$$

If $\lambda, \mu \in \Lambda_1$, it follows that

$$\left(\exp \frac{t}{n} \lambda \exp \frac{t}{n} \mu\right)^n \in H$$

Letting $n \to \infty$ and using the hypothesis that H is closed, we have

$$\lim_{n\to\infty} \left(\exp \frac{t}{n} \lambda \exp \frac{t}{n} \mu\right)^n = \exp t(\lambda + \mu) \in H.$$

for all t. Hence $\lambda + \mu \in \Lambda_1$. To show that Λ_1 is closed under the bracket operation, we may use formula (9) of Lemma 3.8 to obtain the statement

$$\log(\exp t\lambda \exp t\mu \exp -t\lambda \exp -t\mu) = t^2[\lambda, \mu] + O(t^3)$$
$$\text{as } t \to 0.$$

Hence

$$\log\left(\exp \frac{t}{n} \lambda \exp \frac{t}{n} \mu \exp -\frac{t}{n} \lambda \exp -\frac{t}{n} \mu\right) = \log\left(E\left(\frac{t}{n}, \lambda, \mu\right)\right)$$
$$= \frac{t^2[\lambda, \mu]}{n^2} + O\left(\frac{1}{n^3}\right)$$
$$\text{as } n \to \infty$$

If $\lambda, \mu \in H$, it follows exactly as above that

$$\exp t^2[\lambda, \mu] = \lim_{n\to\infty} \left(E\left(\frac{t}{n}, \lambda, \mu\right)\right)^{n^2} \in H.$$

Hence $[\lambda, \mu] \in \Lambda_1$ since $\exp t[\lambda, \mu] \in H$ for all positive t. Thus Λ_1 is a subalgebra of G.

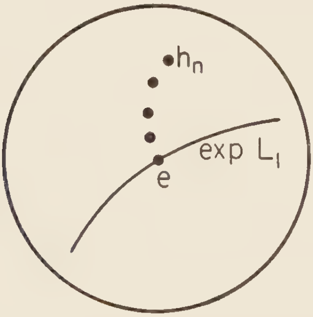
We let Λ_2 be a subspace of Λ which is complementary to $\Lambda_2 : \Lambda = \Lambda_1 + \Lambda_2$ and $\Lambda_1 \cap \Lambda_2 = (0)$. Since $\lambda \to \exp \lambda$ is a local homeomorphism it follows readily that there exist neighborhoods L of 0 in Λ, and L_i of 0 in Λ_i $(i = 1, 2)$ such that the mapping $(\lambda_1, \lambda_2) \to (\exp \lambda_1)(\exp \lambda_2)$ is a homeomorphism of $L_1 \times L_2$ into G. We assume that L is small enough so that the mapping $\lambda \to \exp \lambda$ is a homeomorphism, and we may also assume that $L_i \subseteq L$.

We next show that there is a neighborhood U of 0 in L such that

$$(\exp U) \cap H = \exp U \cap \exp L_1 .$$

To prove this assume the contrary. Then if U_n is any sequence of neighborhoods approaching 0 there exists a sequence of points $h_n \in H$ such that

$$h_n \to e, \text{ yet } h_n \notin \exp L_1 .$$



We write $h_n = (\exp \lambda_{1,n})(\exp \lambda_{2,n})$ where $\lambda_{i,n} \in \Lambda_i$. Hence $\lambda_{i,n} \to 0$ as $n \to \infty$ $(i = 1, 2)$. Multiplying h_n by $\exp(-\lambda_{1,n})$ if necessary, we see that without loss of generality we may assume

$$h_n = \exp \lambda_{2,n}; \quad \lambda_{2,n} \neq 0; \quad \lambda_{2,n} \to 0; \quad \lambda_{2,n} \in \Lambda_2$$

Passing to a subsequence of $\lambda_{2,n}$ if necessary, and using a standard compactness argument we may assume

$$\lambda_{2,n} = a_n \mu_n$$

where the scalars a_n go to 0, where $a_n > 0$, and where $\mu_n \to \mu_0 \in \Lambda_2$ and $\mu_0 \neq 0$. We shall prove that $\mu_0 \in \Lambda_1$, and this will give the desired contradiction since $\Lambda_1 \cap \Lambda_2 = (0)$. For any $t > 0$, write $[t/a_n] = j(n)$, so that

$j(n)$ is an integer, $j(n) \to \infty$, and $a_n j(n) = t + o(1)$. We have

$$\exp a_n \mu_n \in H$$

Hence

$$(\exp a_n \mu_n)^{j(n)} \in H$$

But

$$\log (\exp a_n \mu_n)^{j(n)} = j(n)\, a_n \mu_n \to t\mu_0$$

Hence

$$\exp a_n \mu_n \to \exp t\mu_0$$

Since H is closed, it follows that $\exp t\mu_0 \in H$ for all positive t. Hence $\mu_0 \in \Lambda_1$ which is our contradiction.

We may now proceed to complete the proof of our theorem. According to Theorem 6.3, the subalgebra Λ_1 is the Lie algebra of a connected Lie subgroup H_1 which is generated by elements of the form $\exp \lambda_1$ where λ_1 is in some neighborhood L_1 of 0 in Λ_1. However, we have seen that the set $\exp L_1$ is a neighborhood of e in H (using the subspace topology in H). If we let H_0 be the connected component of the identity in H, and give H_0 the subspace topology, it follows by Lemma 2.2 that H_0 is generated by $\exp L_1$. (Lemma 2.2 was purely topological in statement and proof.) Thus H_1 and H_0 are identical as point sets, and we shall now show that they agree in their topologies. We already know that H_1 has more open sets than H_0; i.e. that any open set of H_0 (using the subspace topology) is open in H_1. (This is because the inclusion map of H_1 in G is continuous.) Conversely, suppose that U is open in H_1. By the definition of the topology of H_1, if $P_0 \in U$, there exists a neighborhood N of e in H_1 such that $P_0 N \subseteq U$. But by what has been proved above, we may write $N = V \cap H$ where V is an open set in G. Thus $(P_0 V) \cap H \subseteq U$, and by the continuity of the group operations, it follows that U is open in H if H has the subspace topology; i.e. U is open in H_0. This shows that the topologies of H_0 and H_1 coincide i.e., that H_1 has the subspace topology. But H is the disjoint union of H_1 and all of its open cosets (Lemma 2.3). Thus H is easily given a C^∞ structure 'by translation' from the C^∞ structure on H_0. Finally, H itself is easily seen to have the subspace topology.

The uniqueness of the C^∞ structure of H is an easy consequence of Theorem 4.14. For if H' is a Lie subgroup of G whose points are those of H, and if H' also has the subspace topology, then the inclusion map $i: H \to H'$ is clearly a continuous homomorphism. Hence, by Theorem 4.14, it is C^∞. The same result applied to the map $i^{-1}: H' \to H$ implies our result. Q.E.D.

Remark. If H is a closed subgroup of G, we shall always give H the C^∞ structure whose existence and uniqueness is shown in the above theorem. It is easily seen that in this case, the tangent space A_1 of H consists of all vectors of the form $\left.\frac{dP}{dt}\right|_{t=0}$, where $P(t)$ is a C^∞ curve of G, $P(0) = e$, and $P(t) \in H$ for all H.

7. Important examples—the classical matrix groups

Much of part II of this book is concerned with the classification of an important special type of Lie Algebra, the so called semi-simple algebras. We shall see that the Lie algebras of the "classical groups" include almost all the possible semi-simple algebras. It therefore is appropriate that we acquire some familiarity with these groups. Since these groups have very simple concrete representations (as groups of matrices), they serve very well to illustrate some of the preceding theory, as well as that which is to come. In what follows, we work strictly with real groups and algebras. For complex algebras and groups analogous results may be stated. The connection between the real and complex cases is developed in part II.

We recall (Lemma 3.15) that $GL(n)$ is the Lie group of all non-singular $n \times n$ matrices and that $A(n)$, the Lie algebra of all $n \times n$ matrices is its Lie algebra. ($[A, B] = AB - BA$.) The first of the classical groups is the subgroup of matrices having determinant 1:

1. Definition. The Special Linear group, $SL(n)$ is the subgroup $\{A | A \in GL(n) \text{ and } \det A = 1\}$ of $GL(n)$.

Remark. $SL(n)$ is an $(n^2 - 1)$ dimensional Lie group. For it is defined by one equation ($\det a_{ij} = 1$) in n^2 variables. $SL(n)$ is clearly closed in $GL(n)$ and it follows from Theorem 6.12 that it is a Lie group. We aim to calculate its Lie algebra—but first, a lemma.

2. Lemma. Let det: $A \to \det(A)$ map $GL(n)$ into the multiplicative group of real numbers, and let $\det_*$ be the corresponding "gradient" map of tangent vectors to $GL(n)$ into the additive group of real numbers. Then

a) $\det_* A = \operatorname{tr} A$.
b) $\exp(\operatorname{tr} A) = \det(\exp A)$.

Proof. Recall that if $A = (a_{ij})$, tr A (the trace of A) is the sum $\sum_{i=1}^{n} a_{ii}$. To calculate $\det_* A$ (evaluated at 1), we use the formula

$$\det(1 + tA) = \begin{vmatrix} 1 + ta_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \\ a_{n1} & \cdots & 1 + ta_{nn} \end{vmatrix}$$

$$= 1 + t(a_{11} + \cdots + a_{nn}) + \cdots$$

$$= 1 + t \operatorname{tr} A + \cdots$$

to first order in t. Hence, using Lemma 1.21,

$$\det_* A = \det_* \frac{d}{dt}(1 + At)\Big|_{t=0} = \frac{d}{dt}\det(1 + At)\Big|_{t=0}$$

$$= \operatorname{tr} A + \cdots\big|_{t=0} = \operatorname{tr} A.$$

This proves part a). To prove b) we note that the map $A \to \det A$ is a homomorphism of $GL(n)$ into $GL(1)$ (the multiplicative group of the reals). Hence, by Lemma 4.9, part b) follows. Q.E.D.

Remark. Similar considerations may be used to prove the familiar result tr (AB) = tr (BA). Indeed, $GL(1)$ is commutative. Hence, since $A \to \det A$ is a homomorphism it follows by Lemma 3.12 that

$$\det_* [A, B] = [\det_* A, \det_* B] = 0$$

or

$$\operatorname{tr}[AB - BA] = 0.$$

3. Theorem. The Lie algebra of $SL(n)$ consists of the set of all $n \times n$ matrices of trace 0.

Proof. If $P(t)$ is any curve $SL(n)$ with $P(0) = 1$, then

$$\operatorname{tr} \frac{dP}{dt}\Big|_{t=0} = \det_* \frac{dP}{dt}\Big|_{t=0} = \frac{d}{dt}\det P\Big|_{t=0} = \frac{d}{dt} 1 = 0$$

Thus, all matrices in $SL(n)$ have trace 0.

Conversely, if $\operatorname{tr} \lambda = 0$, then by Lemma 2

$$\det \exp t\lambda = \exp \operatorname{tr} t\lambda = \exp 0 = 1$$

Hence $\exp t\lambda$ is a curve in $SL(n)$. Plainly, $\left.\frac{d}{dt} \exp t\lambda\right|_{t=0} = \lambda$. Hence $\lambda \in \mathcal{A}(SL(n))$. (See the remark following Theorem 6.12, which justifies treating $\exp t\lambda$ as a curve in $SL(n)$.) Thus, $\mathcal{A}(SL(n))$ consists of all matrices of trace 0, which is the required result. Q.E.D.

The next groups which we consider are the orthogonal groups of various orders.

4. Definition. Let R^n consist of $n \times 1$ matrices (i.e. column vectors), so that the canonical inner product in R^n is given by $(x, y) = x^t y = \sum_{i=1}^{n} x_i y_i$. ($A^t$ is the transpose of the matrix A.) A matrix A is called *orthogonal* if it preserves this inner product: $(Ax, Ay) = (x, y)$. The orthogonal group $O(n)$ is the (closed) group of all orthogonal matrices.

Remark. A is orthogonal if and only if $A^t A = 1$. For, $(Ax, Ay) = (Ax)^t Ay = x^t(A^t A)\, y = x^t y$ if and only if $A^t A = 1$.

5. Theorem. The Lie algebra of the orthogonal group $O(n)$ consists of all skew-symmetric matrices $\lambda(\lambda + \lambda^t = 0)$.

Proof. If $A(s)$ is any curve in $O(n)$ which passes through the identity when $s = 0$, then

$$A(s)^t A(s) = 1, \qquad A(0) = 1. \tag{1}$$

Differentiating equation (1), we have

$$\frac{dA^t}{ds} A + A^t \frac{dA}{ds} = 0.$$

Substituting $s = 0$, and setting $\left.\frac{dA}{ds}\right|_{s=0} = \lambda$, we have

$$\lambda + \lambda^t = 0.$$

Thus all matrices in the Lie Algebra of $O(n)$ are skew-symmetric.

Conversely, suppose $\lambda + \lambda^t = 0$. Then if $A(s) = \exp s\lambda$, we have $\left.\frac{dA}{ds}\right|_{s=0} = \lambda$, and

$$\begin{aligned} A(s)^t\, A(s) &= (\exp s\lambda)^t\, (\exp s\lambda) \\ &= (\exp s\lambda^t)\, (\exp s\lambda) \\ &= (\exp(-s\lambda))\, (\exp s\lambda) = 1 \end{aligned}$$

and $A(t)$ is in $O(n)$.

Thus $\Lambda(O(n))$ consists of all skew-symmetric matrices. Q.E.D.

The final family of classical groups (one for each integer n) are the symplectic groups.

6. Definition. Let J be the following skew-symmetric matrix:

$$J = \begin{bmatrix} \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} & & & 0 \\ & \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} & & \\ & & \ddots & \\ 0 & & & \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} \end{bmatrix}$$

Define $\langle x, y \rangle = x^t J y$. A *symplectic* matrix of order $2n \times 2n$ is a matrix A preserving this inner product: $\langle Ax, Ay \rangle = \langle x, y \rangle$. The symplectic group $S_p(2n)$ is the (closed) group of all symplectic matrices.

Remark. A is a symplectic matrix if and only if $A^t J A = J$. (Cf. the remark following Definition 4.)

7. Theorem. The Lie algebra of $S_p(2n)$ consists of all matrices λ satisfying $\lambda^t J + J\lambda = 0$.

Proof. If $A = A(s)$ is any curve in $S_p(2n)$, then

$$A^t J A = J \qquad A(0) = 1 \tag{2}$$

Differentiating equation (2), we have

$$\frac{dA^t}{ds} JA + A^t J \frac{dA}{ds} = 0$$

Setting $\left.\frac{dA}{ds}\right|_{s=0} = \lambda$ and substituting $s = 0$, we have

$$\lambda^t J + J\lambda = 0 \tag{3}$$

Thus any matrix in the Lie group of $S_p(2n)$ satisfies (3).

Conversely, suppose (3) is satisfied. Then if $A(s) = \exp s\lambda$, then $\left.\frac{dA}{ds}\right|_{s=0} = \lambda$, and

$$A(s)^t JA(s) = (\exp s\lambda^t)\, J(\exp s\lambda) \tag{4}$$

But from the definition of J, we have $J^2 = -1$. Hence, right-multiplying (3) by J, we obtain

$$\lambda^t = J\lambda J \tag{5}$$

Thus, equation (4) implies

$$\begin{aligned} A^t JA &= (\exp s(J\lambda J))\, J(\exp s\lambda) \\ &= (\exp -\, s(J\lambda J^{-1}))\, J \exp s\lambda \\ &= J(\exp -s\lambda)\, J^{-1} J \exp s\lambda = J \end{aligned}$$

and $A(s)$ is in $S_p(2n)$.

Thus, $A(S_p(2n))$ consists of all matrices satisfying equation (3). Q.E.D.

Part II

Complex Semi-Simple Lie Algebras

The theory developed in Part I of the present book shows that connected Lie groups are very well described by their Lie algebras. Now a Lie algebra is a finite-dimensional *linear* algebraic structure; one expects to be able to push the theory of such structures quite far. In Part II, we shall develop the algebraic theory of Lie algebras, ultimately obtaining a complete structure theory for a certain class of these algebras, the so-called semi-simple algebras, and a certain amount of structural information for algebras in general.

Since our methods will be the methods of linear algebra, it will be important for us to work with vector spaces and algebras over the complex field. We will develop the structural analysis first for Lie algebras over the complex field, and then show how the analysis of the real case can be reduced to that of the complex case.

1. Basics

We begin by repeating an earlier definition, generalizing it at the same time to include the complex case.

1. Definition. A real (resp. complex) *Lie algebra* Λ is a vector space over the real (resp. complex) field furnished with a bilinear product $[\lambda, \mu]$ satisfying the identities $[\lambda, \mu] = -[\mu, \lambda]$ and

$$[[\lambda, \mu], \nu] + [[\mu, \nu], \lambda] + [[\nu, \lambda], \mu] = 0 \quad \text{(Jacobi identity).}$$

In what follows, we will always suppose, unless the contrary is explicitly specified, that a Lie algebra is finite dimensional.

In the present Part II the term Lie algebra (unqualified) will be used only in statements true both for real and for complex Lie algebras.

A simple construction leads from real Lie algebras to complex Lie algebras.

2. Definition. Let Λ be a real Lie algebra. Let $\hat{\Lambda}$ be the set of all formal sums $\lambda + i\mu$, $\lambda, \mu \in \Lambda$. We define addition, multiplication by complex numbers, and a product $[\lambda + i\mu, \lambda' + i\mu']$ by the formulae

$$(\lambda + i\mu) + (\lambda' + i\mu') = (\lambda + \lambda') + i(\mu + \mu')$$

$$(x + iy)(\lambda + i\mu) = (x\lambda - y\mu) + i(y\lambda + x\mu)$$

$$[\lambda + i\mu, \lambda' + i\mu'] = ([\lambda, \lambda'] - [\mu, \mu']) + i([\lambda, \mu'] + [\mu, \lambda']).$$

It is easily verified that with these definitions Λ becomes a complex Lie algebra. The complex Lie algebra $\hat{\Lambda}$ is called the *complexification* of the real Lie algebra Λ.

3. Definition: Let V be a real (resp. complex) vector space. A representation ϱ of a real (resp. complex) Lie algebra Λ on V is a mapping which assigns to each $\lambda \in \Lambda$ a linear transformation $\varrho(\lambda)$ on V, and which is such that

(i) $\varrho(\alpha\lambda + \beta\mu) = \alpha\varrho(\lambda) + \beta\varrho(\mu)$ for all $\lambda, \mu \in \Lambda$ and scalars α, β;
(ii) $\varrho[\lambda, \mu] = \varrho(\lambda)\varrho(\mu) - \varrho(\mu)\varrho(\lambda)$.

The set of linear transformations on a vector space V, taken with the product $[A, B] = AB - BA$, is readily seen to form a (possibly infinite dimensional) Lie algebra. Whenever we deal with linear transformations on a vector space, the expression $[A, B]$ will be understood to mean $AB - BA$.

The following definition gives the appropriate generalization of the notion of "eigenvalue" to representations of a Lie algebra.

4. Definition. Let ϱ be a representation of a complex Lie algebra Λ on a vector space V. Then

(a) A linear functional α on Λ is said to be a *weight* of ϱ if there exists a non-zero vector $v \in V$ such that $\varrho(\lambda)\, v = \alpha(\lambda)\, v$ for each $\lambda \in \Lambda$. The vector v is said to be a *weight vector* belonging to the weight α.

(b) If α is a weight of ϱ, a vector v is said to be a *generalized weight vector* belonging to the weight α if there exists an integer N sufficiently large so that $(\varrho(\lambda) - \alpha(\lambda)\, I)^N v = 0$ for all $\lambda \in \Lambda$. The linear space of all generalized weight vectors belonging to the weight α is called the *weight space* of the representation ϱ belonging to the weight α, and will be written V_α.

If α is not a weight of the representation ϱ, we write for convenience $V_\alpha = \{0\}$.

Any linear transformation on a finite dimensional vector space admits an eigenvalue. Our first aim is to generalize this statement to Lie algebras. A few preliminary results are necessary before we can single out the class of Lie algebras for which the generalization is valid.

5. Definition. (a) Let Λ_0 and Λ_1 be linear subspaces of a Lie algebra Λ. Then by $[\Lambda_0, \Lambda_1]$ we mean the linear space generated by all the products $[\lambda_0, \lambda_1]$, $\lambda_0 \in \Lambda$, $\lambda_1 \in \Lambda_1$.

(b) An ideal I of a Lie algebra Λ is a subspace such that $[I, \Lambda] \subseteq I$.

6. Lemma. Let I and J be ideals in a Lie algebra. Then $[I, J]$ is an ideal included in I and J.

Proof. That $[I, J] \subseteq I \cap J$ is obvious. We have $[[\lambda, \mu], \nu] = -[[\mu, \nu], \lambda] - [[\nu, \lambda], \mu] \in [I, J]$ for each $\lambda \in I$, $\mu \in J$, and $\nu \in \Lambda$ by the Jacobi identity. Thus $[[I, J], \Lambda] \subseteq [I, J]$, proving that $[I, J]$ is an ideal. Q.E.D.

7. Definition. (a) If I is an ideal in a Lie algebra, we define the descending sequence of ideals I^n inductively by $I^1 = I$, $I^{n+1} = [I^n, I]$, and the descending sequence of ideals $I^{(n)}$ inductively by $I^{(1)} = I$, $I^{(n+1)} = [I^{(n)}, I^{(n)}]$.

(b) A Lie algebra Λ is *solvable* if $\Lambda^{(n)} = 0$ for n sufficiently large. It is *nilpotent* if $\Lambda^n = 0$ for n sufficiently large.

8. Lemma. (a) Let Λ be a Lie algebra, I an ideal in Λ. Then $I^{(n)} \subseteq I^n$.

(b) A nilpotent Lie algebra is also solvable.

Proof: Statement (a) follows by induction from Definition 7(a); statement (b) is an immediate consequence. Q.E.D.

We now prove that any finite dimensional representation of a complex solvable Lie algebra admits a non-zero weight vector.

9. Theorem. Let Λ be a solvable complex Lie algebra, and ϱ a representation of Λ in a finite dimensional vector space V. Then there exist a vector $v \neq 0$ in V and a weight α of ϱ such that $\varrho(\lambda)\, v = \alpha(\lambda)\, v$ for all $\lambda \in \Lambda$.

Proof. We proceed by induction on the dimension of Λ. If Λ is 1-dimensional, our theorem reduces to the statement that a linear transformation has a non-zero eigenvector.

Next suppose that our assertion is known for any solvable complex algebra whose dimension is less than k. Let the dimension of Λ be k.

If $\Lambda^{(2)} = \Lambda$, it would follow by induction that $\Lambda^{(n)} = \Lambda$ for all n, contradicting the solvability of Λ. Thus $\Lambda^{(2)}$ is a proper subspace of Λ, and there exists a $k - 1$ dimensional subspace M of Λ containing $\Lambda^{(2)}$. Since $[M, M] \subseteq [\Lambda, \Lambda] = \Lambda^{(2)} \subseteq M$, M is a subalgebra of Λ. It follows easily by induction that $M^{(n)} \subseteq \Lambda^{(n)}$; thus M is solvable.

Hence, by inductive hypothesis, there is a weight β defined on M, and a non-zero vector $v_1 \in V$, such that

$$\varrho(\mu)\, v_1 = \beta(\mu)\, v_1 \qquad \mu \in M. \tag{1}$$

Let $\tilde{\lambda}$ be an element of Λ not belonging to M. Since $M \supseteq \Lambda^{(2)}$, we have $[\tilde{\lambda}, \mu] \in M$ if $\mu \in M$. Thus $[\varrho(\tilde{\lambda}), \varrho(\mu)] \in \varrho(M)$ if $\mu \in M$; and therefore for each real t we have $\exp\,(t\,\mathrm{Ad}\,(\varrho(\tilde{\lambda})))\,\varrho(\mu) \in \varrho(M)$ if $\mu \in M$. Using Lemma I.5.3, it follows that for each t, $-\infty < t < \infty$, there exists a $\nu(t, \mu) \in M$ such that

$$\exp\,(-t\varrho\,(\tilde{\lambda}))\,\varrho(\mu)\exp\,(t\varrho\,(\tilde{\lambda})) = \varrho(\nu(t, \mu)).$$

Hence

$$\exp\,(-t\varrho\,(\tilde{\lambda}))\,\varrho(\mu)\exp\,(t\varrho\,(\tilde{\lambda}))\,v_1 = \beta(\nu(t, \mu))\,v_1,$$

i.e.

$$\varrho(\mu)\exp\,(t\varrho\,(\tilde{\lambda}))\,v_1 = \beta(\nu(t, \mu))\exp\,(t\varrho\,(\tilde{\lambda}))\,v_1.$$

This shows that, if $\mu \in M$, the non-zero vector $\exp\,(t\varrho\,(\tilde{\lambda}))\,v_1$ is always an eigenvector of the linear transformation $\varrho(\mu)$. Now the various eigenspaces of $\varrho(\mu)$ are closed sets having no vector but zero in common; hence a continuous curve cannot pass from one eigenspace to another without passing through zero. It follows that $\exp\,(t\varrho\,(\tilde{\lambda}))\,v_1$ belongs to the eigenvalue $\beta(\mu)$ of $\varrho(\mu)$ for all t. Differentiating with respect to t and putting $t = 0$, we find that if V_β is space of vectors defined by the condition (1), then $\varrho(\tilde{\lambda})\,V_\beta \subseteq V_\beta$. Thus the space V_β contains a non-zero eigenvector v of the linear transformation $\varrho(\tilde{\lambda})$, belonging to a certain complex eigenvalue γ. If we define a linear functional α on Λ by putting $\alpha(\tilde{\lambda}) = \gamma$ and $\alpha(\mu) = \beta(\mu)$ for $\mu \in M$, it is clear that $\varrho(\nu)\,v = \alpha(\nu)\,v$ for $\nu \in \Lambda$; thus α is a weight of Λ and v is a non-zero weight vector belonging to this weight.

This completes our inductive step and the theorem follows. Q.E.D.

If T is a linear transformation in a finite dimensional complex vector space V, then V is the direct sum of all the generalized eigenspaces $V_\lambda = \{v | (T - \lambda I)^N v = 0 \text{ for } N \text{ sufficiently large}\}$. The following theorem generalizes this important decomposition to an arbitrary nilpotent complex Lie algebra.

10. Theorem. Let Λ be a nilpotent complex Lie algebra, and ϱ a representation of Λ on a finite dimensional vector space V. Let α_i, $1 \leqslant i \leqslant N$, be the weights of the representation ϱ, and let V_{α_i} be the weight space belonging to the weight α_i. Then each of the spaces V_{α_i} is invariant under each of the transformations $\varrho(\lambda)$, $\lambda \in \Lambda$; and we have the direct sum decomposition

$$V = V_{\alpha_1} \oplus \cdots \oplus V_{\alpha_N}.$$

Proof. We proceed by induction on the dimension of Λ. If this is 1, then our assertion reduces to the known fact that V decomposes into the direct sum of the generalized eigenspaces of any single linear transformation in V.

Next suppose that our assertion is known for any nilpotent complex algebra whose dimension is less than k. Let the dimension of Λ be k. If $\Lambda^2 = \Lambda$, it would follow by induction that $\Lambda^n = \Lambda$ for all n, contradicting the nilpotence of Λ. Thus Λ^2 is a proper subspace of Λ, and there exists a $k - 1$ dimensional subspace M of Λ containing Λ^2. Since $[M, M] \subseteq \Lambda^2 \subseteq M$, M is a subalgebra of Λ. It follows easily by induction that $M^n \subseteq \Lambda^n$; thus M is nilpotent.

Hence, by inductive hypothesis, we have a direct sum decomposition

$$V = V_{\beta_1} \oplus \cdots \oplus V_{\beta_K}, \tag{1}$$

where $\beta_1, \ldots, \beta_k$ are weights defined on M, and where V_{β_i} is the weight space belonging to the weight β_i. Moreover, $\varrho(\mu)\, V_{\beta_i} \subseteq V_{\beta_i}$ for $i = 1, \ldots, K$ and $\mu \in M$.

Let $\tilde{\lambda}$ be an element of Λ not belonging to M. Let $\lambda \in \Lambda$. Then it follows by induction on q that

$$\varrho(\tilde{\lambda})\,(\varrho(\lambda) - \alpha I)^q - (\varrho(\lambda) - \alpha I)^q\, \varrho(\tilde{\lambda}) = -\sum_{i=0}^{q-1} (\varrho(\lambda) - \alpha I)^i\, \varrho([\mathrm{Ad}\,(\lambda)]\,\tilde{\lambda}) \times (\varrho(\lambda) - \alpha I)^{q-i-1}$$

for each complex number α. A second induction shows that there exist integers $C(q, i)$ such that

$$\varrho(\tilde{\lambda})\,(\varrho(\lambda) - \alpha I)^q - (\varrho(\lambda) - \alpha I)^q\, \varrho(\tilde{\lambda}) = \sum_{i=0}^{q-1} C(q, i)\, \varrho([\mathrm{Ad}\,\lambda]^{i+1}\,\tilde{\lambda})\,(\varrho(\lambda) - \alpha I])^{q-i-1}.$$

In this last identity, put $\lambda = \mu$, where $\mu \in M$. Moreover, let v be a vector in one of the spaces V_{β_j}, put $\alpha = \beta_j(\mu)$, let r be such that $(\varrho(\mu) - \beta_j(\mu) I)^r v = 0$, and take $q \geqslant r$. Then we find

$$-(\varrho(\mu) - \beta_j(\mu) I)^q \varrho(\tilde{\lambda}) v = \sum_{i=q-r}^{q-1} C(q, i) \varrho([\mathrm{Ad}(\mu)]^i \tilde{\lambda}) v.$$

Since Λ is nilpotent, $[\mathrm{Ad}(\mu)]^i \tilde{\lambda} = 0$ for i sufficiently large. Thus we find from the last displayed equation that $(\varrho(\mu) - \beta_j(\mu) I)^q \varrho(\tilde{\lambda}) v = 0$ for q sufficiently large.

In this way, we have proved that the spaces V_{β_j} are invariant under the linear transformation $\varrho(\tilde{\lambda})$. Therefore, each of the spaces V_{β_j} decomposes as a direct sum

$$V_{\beta_j} = V_1^{(j)} \oplus \cdots \oplus V_{n_j}^{(j)} \tag{2}$$

of generalized eigenspaces corresponding to the distinct eigenvalues $z_1^{(j)}, \ldots, z_{n_j}^{(j)}$ of the restriction $\varrho(\tilde{\lambda}) | V_{\beta_j}$. Arguing just as in the preceding paragraph, we see that each of the spaces $V_i^{(j)}$, $1 \leqslant i \leqslant n_j$, is invariant under each of the operators $\varrho(\mu)$, $\mu \in M$. Thus, since $\Lambda = M + (\tilde{\lambda})$, each of the spaces $V_i^{(j)}$, $1 \leqslant i \leqslant n_j$, is invariant under each of the operators $\varrho(\lambda)$, $\lambda \in \Lambda$. Define linear functionals $a_i^{(j)}$ on Λ by putting $a_i^{(j)}(\mu) = \beta_j(\mu)$ for $\mu \in M$, $\alpha_i^{(j)}(\tilde{\lambda}) = z_i$, $1 \leqslant i \leqslant n_j$.

We shall now show that no operator $\varrho(\lambda) | V_i^{(j)}$ has any eigenvalue but $\alpha_i^{(j)}(\lambda)$. Suppose that this is false: then there exists an element $\hat{\lambda}$ in Λ such that $\varrho(\hat{\lambda}) | V_i^{(j)}$ has an eigenvalue $\gamma \neq \alpha_i^{(j)}(\hat{\lambda})$. Let $\hat{V}$ be the generalized eigenspace of $\varrho(\hat{\lambda}) | V_i^{(j)}$ corresponding to this eigenvalue. Then, arguing just as in the second paragraph above, we see that $\hat{V}$ is invariant under each of the transformations $\varrho(\lambda)$, $\lambda \in \Lambda$. By Theorem 9, there exists a non-zero vector $\hat{v} \in \hat{V}$ and a linear functional $\hat{a}$ on Λ such that $\varrho(\lambda) \hat{v} = \hat{a}(\lambda) \hat{v}$. Plainly then $\hat{a}(\hat{\lambda}) = \gamma$. But, since $\hat{V} \subseteq V_i^{(j)}$, every vector in $\hat{V}$ is a generalized eigenvector of $\varrho(\lambda)$ belonging to the eigenvalue $\alpha_i^{(j)}(\lambda)$ if either $\lambda \in M$ or $\lambda = \tilde{\lambda}$. Therefore we must have $\alpha_i^{(j)}(\lambda) = \hat{a}(\lambda)$, so $\gamma = \alpha_i^{(j)}(\tilde{\lambda})$, a contradiction. This verifies our assertion that if $\lambda \in \Lambda$, $\varrho(\lambda) | V_i^{(j)}$ has no eigenvalue but $\alpha_i^{(j)}(\lambda)$.

It is then plain that $V_i^{(j)}$ is the weight space corresponding to the weight $\alpha_i^{(j)}(\lambda)$. It follows from (1) and (2) that V is the direct sum of the spaces $V_i^{(j)}$. On the other hand, we have already seen that these spaces are invariant. Thus we have successfully concluded the inductive step of our inductive proof, and the present theorem follows. Q.E.D.

As we observed during the proof of the preceding theorem, if Λ is a nilpotent Lie algebra then Ad (λ) is a nilpotent operator for each $\lambda \in \Lambda$. The following important theorem is just the converse of this statement.

11. Theorem. Let Λ be a Lie algebra, and suppose that Ad (λ) is nilpotent for each $\lambda \in \Lambda$. Then Λ is nilpotent.

We prepare for the proof of Theorem 11 with the following lemma and corollary.

12. Lemma. Let Λ be a Lie algebra of dimension k, and let Ad (λ) be nilpotent for each $\lambda \in \Lambda$. Let ϱ be a representation of Λ in a finite dimensional space V, and suppose that $\varrho(\lambda)$ is nilpotent for each $\lambda \in \Lambda$. Then there exists a non-zero vector $v' \in V$ such that $\varrho(\lambda)\, v' = 0$ for all λ.

13. Corollary. Under the hypotheses of the preceding lemma, V has a descending sequence $V = V_0 \supsetneqq V_1 \supsetneqq V_2 \supsetneqq V_3 \supsetneqq \cdots \supsetneqq V_n = \{0\}$ of subspaces such that $\varrho(\lambda)\, V_i \subseteq V_{i+1}$ for all $\lambda \in \Lambda$.

Proof of Lemma 12. We proceed by induction on the dimension k of Λ. If this is 1, our result is obvious. Suppose then $k > 1$, and that Lemma 12 is known to hold for any proper subalgebra Λ_0 of Λ; we shall show that it holds for Λ.

Since any 1-dimensional subspace of Λ is a subalgebra, Λ surely admits proper subalgebras. Let Λ_0 be a proper subalgebra of Λ of maximal dimension. Then, if $\lambda_0 \in \Lambda_0$, Ad (λ_0) defines a linear transformations of Λ which leaves the subspace $\Lambda_0 \subseteq \Lambda$ invariant. Thus Ad (λ_0) induces a linear transformation of the factor space Λ/Λ_0, which we call $\widehat{\text{Ad}}\,(\lambda_0)$. It is easy to see that $\lambda_0 \to \widehat{\text{Ad}}\,(\lambda_0)$ defines a representation of Λ_0 on the vector space Λ/Λ_0. Moreover, since Ad (λ_0) is nilpotent for each $\lambda_0 \in \Lambda_0$, so is $\widehat{\text{Ad}}\,(\lambda_0)$. Thus, by inductive hypothesis, there exists a nonzero vector $\mu \in \Lambda/\Lambda_0$ such that $\widehat{\text{Ad}}\,(\lambda_0)\,\mu = 0$. That is to say, there exists a vector $\tilde{\lambda} \in \Lambda$, $\tilde{\lambda} \notin \Lambda_0$, such that $[\tilde{\lambda}, \Lambda_0] \subseteq \Lambda_0$.

It is then clear that $(\tilde{\lambda}) + \Lambda_0$ defines a subalgebra of Λ whose dimension exceeds the dimension of Λ_0 by 1. Thus, if Λ_0 is to be a maximal proper subalgebra of Λ, its dimension must be $k - 1$, and we must have $(\tilde{\lambda}) + \Lambda_0 = \Lambda$. By inductive hypothesis, the space $V_0 = \{v | \varrho(\lambda_0)\, v = 0, \lambda_0 \in \Lambda_0\}$ contains non-zero vectors. If $v \in V_0$ and $\lambda_0 \in \Lambda_0$, then $\varrho(\lambda_0)\varrho(\tilde{\lambda})\, v = \varrho(\tilde{\lambda})\, \varrho(\lambda_0)\, v + \varrho([\lambda_0, \tilde{\lambda}])\, v = 0$. This shows that V_0 is invariant under the nilpotent transformation $\varrho(\tilde{\lambda})$, so that V_0 contains a non-zero vector v' such that $\varrho(\tilde{\lambda})\, v' = 0$. Q.E.D.

Proof of Corollary 13. Let V^* be the space of linear functionals on V, and define a representation ϱ^* of Λ on V^* by putting $[\varrho^*(\lambda) v^*](v) = v^*(\varrho(\lambda)v)$ for each $v \in V$. By the preceding lemma, there exists a non-zero $v_0^* \in V^*$ such that $\varrho^*(\lambda) v_0^* = 0$ for $\lambda \in \Lambda$. The space $V_1 = \{v | v_0^*(v) = 0\}$ has then the property $\varrho(\lambda) V \subseteq V_1$ for $\lambda \subset \Lambda$, and an evident induction establishes Corollary 13. Q.E.D.

Proof of Theorem 11. Apply Corollary 13 to the representation $\lambda \to \mathrm{Ad}(\lambda)$ of Λ, to obtain the existence of a descending sequence $\Lambda = \Lambda_1 \supsetneqq \Lambda_2 \supsetneqq \cdots \supsetneqq \Lambda_n = \{0\}$ of subspaces such that $[\Lambda, \Lambda_k] \subseteq \Lambda_{k+1}$. Then plainly $\Lambda^k \subseteq \Lambda_k$, and Λ is nilpotent. Q.E.D.

We conclude this section with a Corollary which will be used later.

14. Corollary. Let Λ be a complex nilpotent Lie algebra, ϱ a representation of Λ in a finite dimensional space V, $\{\alpha\}$ the weights of the representation ϱ, and, for each $\alpha \in \{\alpha\}$, let V_α be the weight space belonging to the weight α. Then, if $\lambda, \mu \in \Lambda$ and $\alpha \in \{\alpha\}$, each vector $v \in V_\alpha$ is a generalized eigenvector of $\varrho(\lambda)\varrho(\mu)$ belonging to the eigenvalue $\alpha(\lambda)\alpha(\mu)$.

Proof. By Theorem 10 each of the spaces V_α is invariant under each of the operators $\varrho(\lambda)$, $\lambda \in \Lambda$. Thus, passing from consideration of the representation ϱ to consideration of the representation $\lambda \to \varrho(\lambda)|V_\alpha$, we may suppose without loss of generality that $V_\alpha = V$. By Theorem 9, there exists a non-zero vector $v_0 \in V$ such that $\varrho(\lambda) v_0 = \alpha(\lambda) v_0$. Thus $\varrho([\lambda, \mu]) v_0 = (\varrho(\lambda)\varrho(\mu) - \varrho(\mu)\varrho(\lambda) v_0 = (\alpha(\lambda)\alpha(\mu) - \alpha(\mu)\alpha(\lambda)) v_0 = 0$. This shows that $\alpha([\lambda, \mu]) = 0$ for all $\lambda, \mu \in \Lambda$. Consequently $\varrho'(\lambda) = \varrho(\lambda) - \alpha(\lambda) I$ defines a representation of Λ in V. Plainly, each of the operators $\varrho'(\lambda)$ is nilpotent in $V = V_\alpha$. Thus, by Corollary 13, so is $\varrho'(\lambda)\varrho'(\mu) + \alpha(\lambda)\varrho'(\mu) + \alpha(\mu)\varrho'(\lambda) = \varrho(\lambda)\varrho(\mu) - \alpha(\lambda)\alpha(\mu) I$, and the present Corollary follows at once. Q.E.D.

2. Cartan's Criterion

1. Definition. (i) A Lie algebra Λ is *simple* if it is of more than one dimension and contains no proper ideals I.

(ii) A Lie algebra Λ is *semi-simple* if it is of more than one dimension and contains no non-zero ideals I such that $I^{(n)} = 0$ for all sufficiently large n.

The reader will easily verify that a simple Lie algebra is semi-simple. In what follows, we shall show that a semi-simple Lie algebra Λ can be decomposed as a direct sum of ideals, each of which is a simple subalgebra of Λ. The methods which we use in proving this first structure theorem will open the way to a complete structural analysis of all simple (and hence all semisimple) Lie algebras.

2. Definition. Let Λ be a Lie algebra. Then the symmetric bilinear form (λ, μ) defined by $(\lambda, \mu) = \operatorname{tr}(\operatorname{Ad}(\lambda)\operatorname{Ad}(\mu))$ is called the *natural bilinear form* of Λ.

Our efforts in the present section will principally be directed to proving the following basic result, known as the "Cartan Criterion".

3. Theorem. A Lie algebra Λ is semi-simple if and only if its natural bilinear form is non-singular.

Before we can prove Theorem 3, however, we must begin to penetrate into the structure of Lie algebras. The following definition prepares the way.

4. Definition. Let Λ be a complex Lie algebra, and M a nilpotent subalgebra of Λ. By the Jacobi identity, the mapping $\mu \to \operatorname{Ad}(\mu)$ is a representation of the Lie algebra M on the vector space Λ. This representation is called the *adjoint representation* of M on Λ. The weights γ of this representation are called the *roots* of M (in Λ); the weight space Λ_γ belonging to the weight γ is called the *root space* belonging to γ.

5. Lemma. Let Λ be a complex Lie algebra, and M a nilpotent subalgebra of Λ. Let γ and γ' be roots of M, and $\Lambda_\gamma, \Lambda_{\gamma'}$ the corresponding root spaces. Then

$$[\Lambda_\gamma, \Lambda_{\gamma'}] \subseteq \Lambda_{\gamma+\gamma'},$$

so that, in particular, $[\Lambda_\gamma, \Lambda_\gamma\ ] = \{0\}$ if $\gamma + \gamma'$ is not a root.

Proof. Using the Jacobi identity, we find by induction that

$$(\operatorname{Ad}(\mu) - (\alpha + \beta) I)^p [\lambda, \nu]$$
$$= \sum_{k=0}^{p} \binom{p}{k} [(\operatorname{Ad}(\mu) - \alpha I)^k \lambda, (\operatorname{Ad}(\mu) - \beta I)^{p-k} \nu]$$

for $\mu \in M$, $\lambda, \nu \in \Lambda$, and all complex numbers α and β. From this identity it is plain that $[\lambda, \nu] \in \Lambda_{\gamma+\gamma'}$ if $\lambda \in \Lambda_\gamma$ and $\nu \in \Lambda_{\gamma'}$, proving our Lemma. Q.E.D.

6. Corollary. If Λ is a complex Lie algebra and M a nilpotent subalgebra of Λ, the root space Λ_0 corresponding to the root 0 is a subalgebra of Λ and includes M.

Proof: By the preceding lemma, $[\Lambda_0, \Lambda_0] \subseteq \Lambda_0$; thus Λ_0 is a subalgebra of Λ. Since M is nilpotent, $[\mathrm{Ad}\,(\mu)]^p\,(\mu') = 0$ for sufficiently large p if $\mu, \mu' \in M$; thus $\Lambda_0 \supseteq M$. Q.E.D.

The following definition singles out a class of subalgebras of a Lie algebra Λ which will be of the greatest use to us in analysing the structure of Λ.

7. Definition. Let Λ be a complex Lie algebra. Let M be a nilpotent subalgebra of Λ, and let Λ_0 be the root space corresponding to the root 0 of M. Then M is a *Cartan subalgebra* if $M = \Lambda_0$.

8. Lemma. Let Λ be a complex Lie algebra, for each $\lambda \in \Lambda$ let $n(\lambda)$ be the number of eigenvalues of the linear transformation $\mathrm{Ad}\,(\lambda)$, and let λ_0 be chosen so that $n(\lambda_0) = \sup_{\lambda \in \Lambda} n(\lambda)$. Then the set M of all $\mu \in \Lambda$ which are generalized eigenvectors of $\mathrm{Ad}\,(\lambda_0)$ belonging to the eigenvalue zero is a Cartan subalgebra of Λ.

Lemma 8 plainly has the following consequence:

9. Corollary. Any complex Lie algebra has a Cartan subalgebra.

Proof of Lemma 8. Let $\hat{\Lambda}_0$ be the one-dimensional subalgebra spanned by λ_0, so that $\hat{\Lambda}_0$ is plainly nilpotent. Then clearly M is the root space belonging to the root 0 of $\hat{\Lambda}_0$; hence, M is subalgebra of Λ by Corollary 6. Let $0, \alpha_1, \ldots, \alpha_k$ be the distinct eigenvalues of $\mathrm{Ad}\,(\lambda_0)$, and let $\varepsilon > 0$ be given. If $\mu \in M$ and the complex number z is sufficiently small, then, by Lemma 0.27, $\mathrm{Ad}\,(\lambda_0 + z\mu)$ has an eigenvalue within a distance ε of each of $\alpha_1, \ldots, \alpha_k$. On the other hand, since M is a subalgebra of Λ, M is invariant under $\mathrm{Ad}\,(\lambda_0 + z\mu)$. Thus, for sufficiently small z all the eigenvalues of the restriction $\mathrm{Ad}\,(\lambda_0 + z\mu)|M$ lie within a distance ε of 0.

Since $(\mathrm{Ad}\,(\lambda_0 + z\mu))\,(\lambda_0 + z\mu) = 0$, zero is always included among these eigenvalues. Thus, if there existed arbitrarily small z for which $\mathrm{Ad}\,(\lambda_0 + z\mu)|M$ had a non-zero eigenvalue, it would follow that for some sufficiently small z $\mathrm{Ad}\,(\lambda_0 + z\mu)$ had more eigenvalues than $\mathrm{Ad}\,(\lambda_0)$, which is impossible. Thus $\mathrm{Ad}\,(\lambda_0 + z\mu)|M$ is nilpotent for all sufficiently small z, i.e. there exists an integer p such that $(\mathrm{Ad}\,(\lambda_0 + z\mu))^p\, M = 0$ for sufficiently small z. This identity must then hold for all complex z; hence $(\mathrm{Ad}\,(\lambda_0) + z\,\mathrm{Ad}\,(\mu))^p\,\mu' = 0$ for $\mu' \in M$ and all z, so that $(\mathrm{Ad}\,(\mu))^p\,\mu'$ is zero for each μ and $\mu' \in M$. By Theorem 1.11, M is nilpotent. Let Λ_0 be the root space corresponding to the root 0 of M. Then $\Lambda_0 \supseteq M$ by Corollary 6. On the other hand, since $\hat{\Lambda}_0 \subseteq M$, Λ_0 is included in the root space belonging to the root 0 of $\hat{\Lambda}_0$. That is $\Lambda_0 \subseteq M$, so that $\Lambda_0 = M$, and M is a Cartan subalgebra of Λ. Q.E.D.

In what follows we will have use for a simple generalization of Lemma 5.

10. Lemma. Let Λ be a complex Lie algebra and M a nilpotent subalgebra of Λ. Let ϱ be a representation of Λ on a vector space V (possibly infinite dimensional). Let ϱ' denote the restriction to M of the representation ϱ, so that ϱ' is a representation of M on V. Let γ be a root of M, and α be a weight of ϱ'; let Λ_γ and V_α be the root space and weight space corresponding to this root and weight respectively. Then if $\lambda \in \Lambda_\gamma$ and $v \in V_\alpha$, we have $\varrho(\lambda)\, v \in V_{\alpha+\gamma}$. In particular, if $\alpha + \gamma$ is not a weight of ϱ', $\varrho(\lambda)\, v = 0$.

Proof. Since ϱ is a representation of Λ, we find by induction that

$$(\varrho(\mu) - (z + Z)\, I)^p \, \varrho(\lambda)\, v = \sum_{k=0}^{p} \binom{p}{k} \varrho((\mathrm{Ad}\,(\mu) - zI)^k \, \lambda)\, (\varrho(\mu) - ZI)^{p-k}\, v$$

for $\mu \in M$, $\lambda \in \Lambda$, $v \in V$, and all complex numbers z and Z. From this identity it is plain that $\varrho(\lambda)\, v \in V_{\alpha+\gamma}$ if $\lambda \in \Lambda_\gamma$ and $v \in V_\alpha$. Q.E.D.

The proof of the following lemma uses a device which, as we shall see later, is the key to the structure theory of semi-simple algebras.

11. Lemma. Let Λ be a complex Lie algebra, M a Cartan subalgebra, and ϱ a representation of M on a finite dimensional vector space V. Suppose that $\Lambda = \Lambda^2$. Then if $\mathrm{tr}\,((\varrho(\mu))^2) = 0$ for all $\mu \in M$, each operator $\varrho(\mu)$, $\mu \in M$, is nilpotent.

Proof. Let $\{\alpha\}$ be the roots of M in Λ, and $\Lambda = \sum_{\alpha \in \{\alpha\}} \oplus \Lambda_\alpha$ the decomposition into root spaces given by Theorem 1.10. By Lemma 5, $[\Lambda_\alpha, \Lambda_\beta] \subseteq \Lambda_{\alpha+\beta}$; since $\Lambda = [\Lambda, \Lambda] = \sum_{\alpha,\beta \in \{\alpha\}} [\Lambda_\alpha, \Lambda_\beta]$, it follows that $\Lambda_0 = \sum_{\substack{\alpha,\beta \in \{\alpha\} \\ \alpha+\beta=0}} [\Lambda_\alpha, \Lambda_\beta] = \sum_{\alpha, -\alpha \in \{\alpha\}} [\Lambda_\alpha, \Lambda_{-\alpha}]$. That is, every vector $\lambda_0 \in M$ is a linear combination of vectors of the form $[\lambda_\alpha, \lambda_{-\alpha}]$, where α, $-\alpha \in \{\alpha\}$, $\lambda_\alpha \in \Lambda_\alpha$, $\lambda_{-\alpha} \in \Lambda_{-\alpha}$.

Suppose then that α, $-\alpha \in \{\alpha\}$, and that $\lambda_\alpha \in \Lambda_\alpha$, $\lambda_{-\alpha} \in \Lambda_{-\alpha}$. Put $\mu = [\lambda_\alpha, \lambda_{-\alpha}]$. Let $\{\eta\}$ be the set of weights of ϱ, and let $V = \sum_{\eta \in \{\eta\}} \oplus V_\eta$ the decomposition into weight spaces given by Theorem 1.10. By Lemma 10,

$\varrho(\lambda_\alpha)\, v \in V_{\eta+\alpha}$ if $v \in V_\eta$, while $\varrho(\lambda_{-\alpha})\, v \in V_{\eta-\alpha}$. Thus the subspace

$$V' = \sum_{-\infty<k<+\infty} \oplus\, V_{\eta+k\alpha}$$

of V is invariant under $\varrho(\lambda_\alpha)$ and $\varrho(\lambda_{-\alpha})$. Hence

$$\begin{aligned} \operatorname{tr}(\varrho(\mu)|V') &= \operatorname{tr}(\varrho\,[(\lambda_\alpha, \lambda_{-\alpha}])|V') = \operatorname{tr}(\{\varrho(\lambda_\alpha)\,\varrho(\lambda_{-\alpha}) - \varrho(\lambda_{-\alpha})\,\varrho(\lambda_\alpha)\}|V') \\ &= \operatorname{tr}([\varrho\,(\lambda_\alpha)|V']\,[\varrho(\lambda_{-\alpha})|V'] - [\varrho(\lambda_{-\alpha})|V']\,[\varrho(\lambda_\alpha)|V']) = 0. \end{aligned}$$

On the other hand, since $(\varrho(\mu) - (\eta(\mu) + k\alpha(\mu))\, I)^p\, V_{\eta+k\lambda} = 0$ for p sufficiently large, we have $\operatorname{tr}(\varrho(\mu)|V_{\eta+k\alpha}) = (\eta(\mu) + k\alpha(\mu)) \dim(V_{\eta+k\alpha})$. Thus

$$\begin{aligned} 0 = \operatorname{tr}(\varrho(\mu)|V') &= \sum_{-\infty<k<+\infty} \operatorname{tr}(\varrho(\mu)|V_{\eta+k\alpha}) \\ &= \sum_{-\infty<k<+\infty} (\eta(\mu) + k\alpha(\mu)) \dim(V_{\eta+k\alpha}). \end{aligned}$$

It follows that for each weight η

$$\eta(\mu) = -\left(\frac{\sum k \dim(V_{\eta+k\alpha})}{\sum \dim(V_{\eta+k\alpha})}\right)\alpha(\mu).$$

That is, $\eta(\mu) = r_{\eta,\alpha}\alpha(\mu)$ where $r_{\eta,\alpha}$ is *rational.* Now, by Corollary 1.14, each vector in V_η is a generalized eigenvector of $(\varrho(\mu))^2$ belonging to the eigenvalue $(\eta(\mu))^2$. Thus

$$0 = \operatorname{tr}((\varrho(\mu))^2) = \sum_{\eta\in\{\eta\}} (\eta(\mu))^2 \dim(V_\eta) = \sum_{\eta\in\{\eta\}} (r_{\eta,\alpha})^2\, (\alpha(\mu))^2 \dim(V_\eta).$$

This makes it plain that every term in the sum $\sum_{\eta\in\{\eta\}} (\eta(\mu))^2 \dim(V_\eta)$ must be zero. Thus $\eta(\mu) = 0$ for $\mu \in [\Lambda_\alpha, \Lambda_{-\alpha}]$. Since $M = \Lambda_0 = \sum [\Lambda_\alpha, \Lambda_{-\alpha}]$, all the eigenvalues $\eta(\lambda_0) = 0$ of $\varrho(\lambda_0)$ are zero for each $\lambda_0 \in M$, and thus $\varrho(\lambda_0)$ is nilpotent. Q.E.D.

We record an identity from the preceding proof for future use.

12. Corollary. Let Λ be a complex Lie algebra, M a Cartan subalgebra of Λ, and ϱ a representation of M on a finite dimensional vector space V. Suppose that $\Lambda = \Lambda^2$. Let $\{\eta\}$ be the set of weights of ϱ, and let $\{\alpha\}$ be the set of roots of M in Λ. Let V_η and Λ_α be the corresponding weight spaces and root spaces. Let $\mu \in [\Lambda_\alpha, \Lambda_{-\alpha}]$. Then $\eta(\mu) = r_{\eta,\alpha}\alpha(\mu)$, where the number $r_{\eta,\alpha}$ is rational, and is given by

$$r_{\eta,\alpha} = -\frac{\sum\limits_{-\infty<k<+\infty} k \dim(V_{\eta+k\alpha})}{\sum\limits_{-\infty<k<+\infty} \dim(V_{\eta+k\alpha})}.$$

We are now in a position to prove Cartan's Criterion:

Proof of Theorem 3. First suppose that the algebra Λ is complex. Let (λ, μ) be its natural bilinear form. If this form is not non-singular, then $\Lambda_0 = \{\lambda | (\lambda, \mu) = 0 \text{ for all } \mu \in \Lambda\}$ defines a non-trivial subspace of Λ. Since $(\mathrm{Ad}(\nu)\lambda, \mu) = -\mathrm{tr}(\mathrm{Ad}([\lambda, \nu])\mathrm{Ad}(\mu)) = -\mathrm{tr}([\mathrm{Ad}(\lambda), \mathrm{Ad}(\nu)]\mathrm{Ad}(\mu)) = -\mathrm{tr}(\mathrm{Ad}(\lambda)\{\mathrm{Ad}(\nu)\mathrm{Ad}(\mu) - \mathrm{Ad}(\mu)\mathrm{Ad}(\nu)\}) = -(\lambda, \mathrm{Ad}(\nu)\mu)$, for $\nu, \lambda, \mu \in \Lambda$, it follows that $\mathrm{Ad}(\lambda)(\Lambda_0) \subseteq \Lambda_0$. Thus $[\Lambda, \Lambda_0] \subseteq \Lambda_0$, i.e., Λ_0 is an ideal of Λ. We shall show that Λ_0 is solvable.

Suppose that this is false, i.e., that Λ_0 is not solvable. Then the decreasing sequence $\Lambda_0 \supseteq \Lambda_0^{(2)} \supseteq \Lambda_0^{(3)} \supseteq \cdots$ of ideals eventually reaches a non-zero ideal $\Lambda_1 = \Lambda_0^{(k)}$ such that $\Lambda_1 = \Lambda_1^{(2)} = \Lambda_1^2$. Regard the subalgebra Λ_1 of Λ as a Lie algebra in its own right, and let $M \subseteq \Lambda_1$ be a Cartan subalgebra of Λ_1. Then $\mu \to \mathrm{Ad}(\mu)$ defines a representation of M on Λ; since $M \subseteq \Lambda_0$, we have $\mathrm{tr}((\mathrm{Ad}(\mu))^2) = 0$, i.e., $(\mu, \mu) = 0$ for $\mu \in M$. Thus, by Lemma 11, $\mathrm{Ad}(\mu)$ is nilpotent for each $\mu \in M$. But then by the definition of root spaces each $\lambda_1 \in \Lambda_1$ belongs to the root space of M in Λ_1 corresponding to the root $\alpha = 0$. Thus, by Definition 7, $M = \Lambda_1$; thus Λ_1 is nilpotent, and $\Lambda_1 = \Lambda_1^2$ is impossible. Thus we have proved that if Λ is a complex semi-simple algebra, the natural bilinear form of Λ is non-singular.

Conversely suppose that Λ has a solvable ideal M. Then, for n sufficiently large, $M^{(n)} = P$ will be a non-zero ideal such that $P^{(2)} = 0$. For $\mu \in P$ we therefore have $\mathrm{Ad}(\mu)\Lambda \subseteq P$, $\mathrm{Ad}(\mu)P = \{0\}$. Consequently $(\mathrm{Ad}(\mu)\mathrm{Ad}(\lambda))^2 = 0$ for $\mu \in P$ and $\lambda \in \Lambda$, so that $(\mu, \lambda) = tr(\mathrm{Ad}(\mu)\mathrm{Ad}(\lambda)) = 0$ for all $\mu \in P$ and $\lambda \in \Lambda$, so that the natural bilinear form of Λ is singular.

Next we consider the case Λ real. Let $\hat{\Lambda}$ be the complexification of Λ. Let (λ, λ') be the natural bilinear form of Λ, and $(\mu, \mu')_c$ be that of $\hat{\Lambda}$ then it is easily seen that for each $\lambda, \nu, \lambda', \nu' \in \Lambda$ we have

$$(\lambda + i\nu, \lambda' + i\nu')c = (\lambda, \lambda') - (\nu, \nu') + i(\lambda, \nu') + i(\lambda', \nu). \tag{1}$$

Thus Λ has a non-singular natural bilinear form if and only if $\hat{\Lambda}$ does.

For each ideal M of Λ, put $\hat{M} = \{\lambda + i\nu \,|\, \lambda, \nu \in M\}$, so that $\hat{M}$ is an ideal of $\hat{\Lambda}$. It follows readily by induction that $\widehat{M^{(n)}} = \hat{M}^{(n)}$; thus if Λ has a solvable ideal, so does $\hat{\Lambda}$. Suppose conversely that $\hat{\Lambda}$ has a solvable ideal. Then by what has already been proved for complex Lie algebras (cf. the final statement of the first paragraph of the present proof), $\Lambda_0' = \{\mu \in \hat{\Lambda} | (\mu, \mu')_c = 0, \mu' \in \hat{\Lambda}\}$ is a solvable ideal of $\hat{\Lambda}$. It is plain from (1) that if $(\lambda + i\nu) \in \Lambda_0'$, then $\lambda + i\,0$ and $\nu + i\,0$ both belong to Λ_0'. Thus, if we let $\Lambda_0 = \{\lambda \in \Lambda |\, \lambda + i\nu \in \Lambda_0' \text{ for some } \nu\}$, we have $\Lambda_0' = \hat{\Lambda}_0$, and therefore, as we have already remarked, $\widehat{\Lambda_0^{(n)}} = (\hat{\Lambda}_0)^{(n)}$, so that Λ_0 is solvable. This shows that Λ has a solvable ideal if and only if $\hat{\Lambda}$ has a solvable ideal. The real case of Theorem 3 now follows immediately from its complex case, which we have already proved. Q.E.D.

13. Corollary. A real Lie algebra Λ is semi-simple if and only if its complexification $\hat{\Lambda}$ is semi-simple.

14. Corollary. Let Λ be a semi-simple Lie algebra, and (λ, μ) its natural bilinear form. Then Λ contains no non-zero ideal M such that $(\lambda, \mu) = 0$ for all $\lambda, \mu \in M$.

Proof. Suppose that such an ideal M existed. First let Λ be complex. Then if in the second paragraph of the preceding proof we replace Λ_0 by M, the argument of that paragraph would show that M is solvable, contradicting the semi-simplicity of Λ. This proves our assertion if Λ is complex.

Next let Λ be real, and $\hat{\Lambda}$ its complexification. Define $\hat{M}$ as in the final paragraph of the preceding proof. Then, by formula (1) of that proof, $(\mu, \mu')_c = 0$ for $\mu, \mu' \in \hat{M}$. Thus, by what we have already proved $\hat{M} = \{0\}$, and hence $M = \{0\}$. Q.E.D.

15. Corollary. The natural bilinear form of a Lie algebra Λ satisfies the identity $(\mathrm{Ad}\,(\lambda)\,\mu, \nu) = -(\mu, \mathrm{Ad}\,(\lambda)\,\nu)$, $\lambda, \mu, \nu \in \Lambda$.

The Cartan Criterion has the following important structural theorem as an elementary consequence.

16. Theorem. A Lie algebra Λ is semi-simple if and only if Λ may be written as a direct sum $\Lambda = \Lambda_1 \oplus \cdots \oplus \Lambda_n$, where each Λ_i is an ideal of Λ which, regarded as a Lie algebra in its own right, is simple.

Proof. First suppose that Λ is semi-simple. If Λ is simple, we have nothing to prove. If not, Λ contains an ideal Λ'. Let k be the dimension of Λ, and j the dimension of Λ', so that

$$\Lambda'' = \{\lambda \in \Lambda |\, (\lambda, \lambda') = 0 \text{ for } \lambda' \in \Lambda'\}$$

is $k - j$ dimensional; here, of course, (λ, μ) is the natural bilinear form of Λ. Since $(\mathrm{Ad}(\nu)\lambda, \mu) = -(\lambda, \mathrm{Ad}(\nu)\mu)$ for $\lambda, \mu, \nu \in \Lambda$ by Corollary 15, we see easily that Λ'' is an ideal. Thus $\Lambda' \cap \Lambda'' = \Lambda'''$ is an ideal. But if $\mu \in \Lambda'''$, it is plain that $(\mu, \mu) = 0$. By Corollary 14, $\Lambda''' = 0$. Hence $\Lambda = \Lambda' \oplus \Lambda''$; that is, Λ is the direct sum of two ideals. Regard the subalgebras Λ' and Λ'' of Λ as Lie algebras in their own right. If either, say Λ', has an non-zero ideal P such that $P^{(n)} = 0$ for sufficiently large n, then P is also a solvable ideal of Λ, contradicting the assumed semi-simplicity of Λ. Thus both Λ' and Λ'' are semi-simple algebras, and now an evident induction gives the asserted direct sum decomposition of Λ into simple ideals.

Suppose, conversely, that $\Lambda = \Lambda_1 \oplus \cdots \oplus \Lambda_n$, where each Λ_i is an ideal of Λ, which, regarded as a Lie algebra in its own right, is simple. Let $(\lambda, \mu)_i$ be the natural bilinear form of the algebra Λ_i, and (λ, μ) that of Λ. Then if $\lambda_i, \mu_i \in \Lambda_i$, it is easily seen that

$$(\sum \lambda_i, \sum \mu_i) = \sum (\lambda_i, \mu_i)_i.$$

By Theorem 3, all the forms $(\lambda, \mu)_i$ are non-singular; hence by the above formula (λ, μ) is non-singular, and applying Theorem 3 once more, we see that Λ is semi-simple. Q.E.D.

17. Corollary. If Λ is a semi-simple algebra, $\Lambda^2 = \Lambda$.

Proof. Use the decomposition given by the preceding theorem, to obtain $\Lambda^2 = \Lambda_1^2 \oplus \Lambda_2^2 \oplus \cdots \oplus \Lambda_n^2$. Since each Λ_i is simple, its ideal Λ_i^2 cannot be proper, so we must have $\Lambda_i^2 = \Lambda_i$. Q.E.D.

3. Structure of Simple Complex Algebras

The road to a complete structural analysis of all simple (and hence all semi-simple) complex algebras now lies open before us.

Throughout the present section we let Λ be a semi-simple complex Lie algebra, M a Cartan subalgebra of Λ, $\{\alpha\}$ the set of roots of M in Λ, Λ_α the corresponding root spaces, (and $\Lambda_\alpha = \{0\}$ if α is not a root), and let (λ, μ) *be the natural bilinear form of Λ*, so that by Theorem 2.3 (λ, μ) is non-singular.

1. Lemma. If $\lambda_\alpha \in \Lambda_\alpha$, $\lambda_\beta \in \Lambda_\beta$, and $\alpha + \beta \neq 0$, then $\mathrm{Ad}(\lambda_\alpha)\,\mathrm{Ad}(\lambda_\beta)$ is nilpotent, so that $(\lambda_\alpha, \lambda_\beta) = 0$.

Proof. By Lemma 2.5, $[\mathrm{Ad}(\lambda_\alpha)\,\mathrm{Ad}\,(\lambda_\beta)]^n$ maps Λ_γ into $\Lambda_{\gamma+n(\alpha+\beta)}$. If $\alpha + \beta \neq 0$, not all $\gamma + n(\alpha + \beta)$ can be roots, and hence a sufficiently high power of $\mathrm{Ad}\,(\lambda_\alpha)\,\mathrm{Ad}\,(\lambda_\beta)$ maps Λ_γ into zero. Since $\Lambda = \sum_{\gamma \in \{\alpha\}} \oplus \Lambda_\gamma$, $\mathrm{Ad}\,(\lambda_\alpha)\,\mathrm{Ad}\,(\lambda_\beta)$ is nilpotent, and thus its trace $(\lambda_\alpha, \lambda_\beta)$ is zero. Q.E.D.

2. Corollary. If α is a root, so is $-\alpha$. If $\lambda_\alpha \in \Lambda_\alpha$, there exists a $\lambda_{-\alpha} \in \Lambda_{-\alpha}$ such that $(\lambda_\alpha, \lambda_{-\alpha}) = 1$.

Proof. If not, then, by the preceding lemma, $(\lambda_\alpha, \lambda_\gamma) = 0$ for $\lambda_\alpha \in \Lambda_\alpha$ and $\lambda_\gamma \in \Lambda_\gamma$, no matter what γ. But since $\Lambda = \sum_{\gamma \in \{\alpha\}} \oplus \Lambda_\gamma$, the natural bilinear form would then be singular. Q.E.D.

3. Corollary. The restriction of (λ, μ) to $M = \Lambda_0$ is non-singular.

Proof. If $\exists \mu$, $0 \neq \mu \in M$ such that $(\mu, \lambda) = 0$ for all $\lambda \in \Lambda_0$, then by Lemma 1 we would have $(\mu, \lambda) = 0$ for all $\lambda \in \Lambda_\gamma$, no matter what γ. Thus, since $\Lambda = \sum_{\gamma \in \{\alpha\}} \oplus \Lambda_\gamma$, it would follow that $(\mu, \lambda) = 0$ for all $\lambda \in \Lambda$, and the natural bilinear form would then be singular. Q.E.D.

4. Lemma. If $\mu \in M$ and $\alpha(\mu) = 0$ for all α, then $\mu = 0$.

Proof. The spaces Λ_β are invariant under all the maps $\mathrm{Ad}\,(\mu)$, $\mu \in M$. By Theorem 1.9 (and since the nilpotent algebra M is also solvable) there exists a non-zero vector $\lambda \in \Lambda_\beta$ and a linear functional γ on M such that $(\mathrm{Ad}\,(\mu) - \gamma(\mu)\,I)\,\lambda = 0$, $\mu \in M$. Since each $\lambda \in \Lambda_\beta$ is a generalized eigenvector of $\mathrm{Ad}\,(\mu)$ belonging to the eigenvalue $\beta(\mu)$, we must plainly have $\gamma = \beta$. Thus $\mathrm{Ad}\,(\mu)\,\lambda = \beta(\mu)\,\lambda$ for $\mu \in M$. Consequently,

$$\beta([\mu, \mu'])\,\lambda = \mathrm{Ad}\,([\mu, \mu'])\,\lambda = (\mathrm{Ad}\,(\mu)\,\mathrm{Ad}\,(\mu') - \mathrm{Ad}\,(\mu')\,\mathrm{Ad}\,(\mu))\,\lambda$$

$$= (\beta(\mu)\,\beta(\mu') - \beta(\mu')\,\beta(\mu))\,\lambda = 0\,,$$

i.e. $\beta([\mu, \mu']) = 0$ for all $\mu, \mu' \in M$. Using this identity, the reader will verify at once that $\varrho(\mu) = (\mathrm{Ad}\,(\mu) - \beta(\mu)\,I)|\Lambda_\beta$ defines a representation of M on Λ_β, and that each operator $\varrho(\mu)$ is nilpotent. Thus, by Corollary 1.13, $\varrho(\mu)\,\varrho(\mu') + \beta(\mu)\,\varrho(\mu') + \beta(\mu')\,\varrho(\mu)$ is nilpotent for $\mu, \mu' \in M$, i.e., $(\mathrm{Ad}\,(\mu)\,\mathrm{Ad}\,(\mu') - \beta(\mu)\,\beta(\mu')\,I)|\Lambda_\beta$ is nilpotent. Hence using $\Lambda = \sum_{\alpha \in \{\alpha\}} \oplus \Lambda_\alpha$ we have

$$(\mu, \mu') = \mathrm{tr}\,(\mathrm{Ad}\,(\mu)\,\mathrm{Ad}\,(\mu')) = \sum_{\alpha \in \{\alpha\}} \alpha(\mu)\,\alpha(\mu')\,\dim\,(\Lambda_\alpha),\ \mu, \mu' \in M$$

If $\alpha(\mu) = 0$, we have consequently $(\mu, \mu') = 0$ for all $\mu' \in M$, and thus $\mu = 0$ by Corollary 3. Q.E.D.

5. Corollary. $[\mu, \mu'] = 0$ if $\mu, \mu' \in M$.

Proof. We saw in the course of proving the preceding lemma that $\alpha([\mu, \mu']) = 0$ if $\alpha \in \{\alpha\}$. But then, by the statement of the lemma, $[\mu, \mu'] = 0$. Q.E.D.

The following Lemma is a special case of Corollary 2.12.

6. Lemma. If $\mu \in [\Lambda_\alpha, \Lambda_{-\alpha}]$ we have

$$\beta(\mu) = -\frac{\sum\limits_{-\infty < k < +\infty} k \dim(\Lambda_{\beta + k\alpha})}{\sum\limits_{-\infty < k < +\infty} \dim(\Lambda_{\beta + k\alpha})}\alpha(\mu).$$

7. Lemma. If $\alpha \neq 0$ is a root, Λ_α is 1-dimensional, while $\Lambda_{2\alpha} = \Lambda_{3\alpha} = \cdots = 0$.

Proof. Let $\lambda_\alpha \in \Lambda_\alpha$, and, using Corollary 2, pick $\lambda_{-\alpha} \in \Lambda_{-\alpha}$ so that $(\lambda_\alpha, \lambda_{-\alpha}) = 1$. Arguing as in the first paragraph of Lemma 4, we may also suppose that $[M, \lambda_{-\alpha}] \subseteq (\lambda_{-\alpha})$. Let V be the subspace of Λ defined by $V = (\lambda_{-\alpha}) \oplus \sum\limits_{r>0} \oplus \Lambda_{r\alpha}$, r an integer.

By Lemma 2.5, V is invariant under the mappings Ad (λ_α) and Ad $(\lambda_{-\alpha})$. Put $\mu = [\lambda_\alpha, \lambda_{-\alpha}]$. Then

$$\begin{aligned} \operatorname{tr}(\operatorname{Ad}(\mu)|V) &= \operatorname{tr}([\operatorname{Ad}(\lambda_\alpha)|V][\operatorname{Ad}(\lambda_{-\alpha})|V] - \operatorname{tr}[\operatorname{Ad}(\lambda_{-\alpha})|V][\operatorname{Ad}(\lambda_\alpha)|V]) \\ &= 0. \end{aligned}$$

On the other hand, since $(\operatorname{Ad}(\mu) - r\alpha(\mu)\, I)^P|\Lambda_{r\alpha} = 0$ for P sufficiently large, we have $\operatorname{tr}(\operatorname{Ad}(\mu)|\Lambda_{r\alpha}) = r\alpha(\mu) \cdot \dim(\Lambda_{r\alpha})$. Thus

$$\begin{aligned} 0 = \operatorname{tr}(\operatorname{Ad}(\mu)|V) &= -\alpha(\mu) + \sum_{r\geq 0} r\alpha(\mu) \cdot \dim(\Lambda_{r\alpha}) \\ &= \alpha(\mu)\left(-1 + \sum_{r\geq 0} r \dim(\Lambda_{r\alpha})\right). \end{aligned}$$

It is clear from this identity that if the present lemma is false we must have $\alpha(\mu) = 0$. But then $\beta(\mu) = 0$ for all $\beta \in \{\alpha\}$ by Lemma 6, so that by Lemma 4, $\mu = 0$. This would give

$$\begin{aligned} 0 = (\nu, \operatorname{Ad}(\lambda_\alpha)\,\lambda_{-\alpha}) &= -(\operatorname{Ad}(\lambda_\alpha)\,\nu, \lambda_{-\alpha}) \\ &= (\operatorname{Ad}(\nu)\,\lambda_\alpha, \lambda_{-\alpha}) = \alpha(\nu)\,(\lambda_\alpha, \lambda_{-\alpha}) = \alpha(\nu) \end{aligned}$$

for all $\nu \in M$, and hence $\alpha = 0$, contradicting our hypothesis. This contradiction proves our Lemma. Q.E.D.

8. Corollary. If $\lambda_\alpha \in \Lambda_\alpha$ and $\lambda_{-\alpha} \in \Lambda_{-\alpha}$, are such that $(\lambda_\alpha, \lambda_{-\alpha}) = 1$, then the element $\mu_\alpha = [\lambda_\alpha, \lambda_{-\alpha}]$ is non-zero and belongs to M. We have $(\nu, \mu_\alpha) = \alpha(\nu)$ for each $\nu \in M$. Thus μ_α is defined by α, independently of the choice of λ_α and $\lambda_{-\alpha}$. We have $\mu_{-\alpha} = -\mu_\alpha$, $\mu_{\alpha+\beta} = \mu_\alpha + \mu_\beta$.

Proof. By the last part of the preceding argument, $[\lambda_\alpha, \lambda_{-\alpha}] \neq 0$. Then

$$\begin{aligned} -\alpha(\nu)(\lambda_\alpha, \lambda_{-\alpha}) &= -(\mathrm{Ad}\,(\nu)\,\lambda_\alpha, \lambda_{-\alpha}) = (\mathrm{Ad}\,(\lambda_\alpha)\,\nu, \lambda_{-\alpha}) \\ &= -(\nu, \mathrm{Ad}\,(\lambda_\alpha)\,\lambda_{-\alpha}) = -(\nu, [\lambda_\alpha, \lambda_{-\alpha}]), \nu \in M. \end{aligned}$$

Thus $\alpha(\nu) = (\nu, \mu_\alpha)$, $\nu \in M$. The two final statements follow at once from this equation and from Corollary 3. Q.E.D.

Throughout the remainder of the present section, whenever $\alpha \in \{\alpha\}$, *we shall let* μ_α *denote the element of M described in Corollary* 8.

9. Corollary. If $\lambda_\alpha \in \Lambda_\alpha$ and $\mu \in M$, then $[\mu, \lambda_\alpha] = \alpha(\mu)\,\lambda_\alpha$.

Proof. Every vector in the one-dimensional space Λ_α is a generalized eigenvector of Ad (μ) belonging to the eigenvalue $\alpha(\mu)$. Q.E.D.

10. Corollary. The vectors μ_α, $\alpha \in \{\alpha\}$, span M.

Proof. If $\nu \in M$ and $(\nu, \mu_\alpha) = 0$ for $\alpha \in \{\alpha\}$ then $\alpha(\nu) = 0$ for $\alpha \in \{\alpha\}$ by Corollary 8, so that $\nu = 0$ by Lemma 4. Let m be the dimension of M. If all the vectors μ_α were contained in a proper subspace of M, then the conditions $(\nu, \mu_\alpha) = 0$ would be implied by less than m linearly independent conditions, and there would exist a non-zero vector ν satisfying $(\nu, \mu_\alpha) = 0$, $\alpha \in \{\alpha\}$, a contradiction. Q.E.D.

11. Lemma. Let $\alpha, \beta \in \{\alpha\}$, and let $\lambda_\alpha \in \Lambda_\alpha$, $\lambda_\beta \in \Lambda_\beta$ be non-zero. Then, if $\alpha + \beta$ is a root, $[\lambda_\alpha, \lambda_\beta]$ is a non-zero vector of $\Lambda_{\alpha+\beta}$.

Proof: Lemma 2.5 shows that all we have to prove is that $[\lambda_\alpha, \lambda_\beta]$ is non-zero if $\alpha + \beta$ is a root. We prove this as follows: For each $\gamma \in \{\alpha\}$, let λ_γ be a non-zero vector of Λ_γ, chosen so that we always have $(\lambda_\gamma, \lambda_{-\gamma}) = 1$, so that by Corollary 8 $\mu_\gamma = [\lambda_{-\gamma}, \lambda_\gamma]$. Let R be the set of integers r for which $\alpha + r\beta$ is a root. An interval $r_- \leqslant r \leqslant r_+$ of integers is said to be *isolated* if $r_- \in R$, $r_+ \in R$, $[\lambda_{-\beta}, \lambda_{\alpha+r_-\beta}] = 0$ and $[\lambda_\beta, \lambda_{\alpha+r_+\beta}] = 0$. If such an isolated interval is given, then, by Lemma 2.5, the space

$$V = \sum_{r_- \leq r \leq r_+} \oplus \Lambda_{\alpha+r\beta}$$

is invariant under Ad (λ_β) and Ad $(\lambda_{-\beta})$. Thus

$$\operatorname{tr}(\operatorname{Ad}(\mu_\beta)|V) = \operatorname{tr}([\operatorname{Ad}(\lambda_{-\beta})|V][\operatorname{Ad}(\lambda_\beta)|V] - [\operatorname{Ad}(\lambda_\beta)|V][\operatorname{Ad}(\lambda_{-\beta})|V)] = 0.$$

On the other hand, by Lemma 7 and Corollary 8,

$$\operatorname{tr}(\operatorname{Ad}(\mu_\beta)|V) = \sum_{\substack{r_- \leq r \leq r_+ \\ r \in R}} (\alpha(\mu_\beta) + r\beta(\mu_\beta)).$$

This shows that the ratio $-\frac{\alpha(\mu_\beta)}{\beta(\mu_\beta)}$ is equal to the average value A_{r_-,r_+} of the integer r over the set $\{r \in R | r_- \leqslant r \leqslant r_+\}$; here, $r_- \leqslant r \leqslant r_+$ can be *any* isolated interval. If there were any $r^* \in R$ such that $r^* > r_+$, we could find a second isolated interval $r_- \leqslant r \leqslant r_{++}$, with $r_{++} > r_+$. But it is plain that the average value of r over this second interval would be strictly larger than A_{r_-,r_+}, which is a contradiction. A similar argument applies to r_-; so that we see that there can be only one isolated interval, i.e., that R itself must be an interval of the form $r_- \leqslant r \leqslant r_+$, and that $[\lambda_{-\beta}, \lambda_{\alpha+r\beta}] \neq 0$ if $r \in R$ and $r > r_-$, $[\lambda_\beta, \lambda_{\alpha+r\beta}] \neq 0$ if $r \in R$ and $r < r_+$. Our lemma is the special case $r = 0$ of this last assertion. Q.E.D.

12. Corollary. Let $\alpha, \beta \in \{\alpha\}$. The set R of integers for which $\alpha + r\beta \in \{\alpha\}$ is an interval of the form $r_- \leqslant r \leqslant r_+$, and we have

$$-\frac{\alpha(\mu_\beta)}{\beta(\mu_\beta)} = \frac{1}{2}(r_+ + r_-).$$

13. Lemma. If $\alpha, \beta \in \{\alpha\}$, $\beta(\mu_\alpha) = (\mu_\alpha, \mu_\beta)$ is real.

Proof. Using Lemma 6 and Corollary 8 we may write $\beta(\mu_\alpha) = r_{\alpha,\beta}(\mu_\alpha, \mu_\alpha)$ where $r_{\alpha,\beta}$ is rational. Since $\Lambda = \sum_{\beta \in \{\alpha\}} \oplus \Lambda_\beta$ and each vector in Λ_β belongs to the eigenvalue $\beta(\mu_\alpha)$ of Ad (μ_α), we have

$$\begin{aligned}(\mu_\alpha, \mu_\alpha) = \operatorname{tr}(\operatorname{Ad}(\mu_\alpha)\operatorname{Ad}(\mu_\alpha)) &= \sum_{\beta \in \{\alpha\}} (\beta(\mu_\alpha))^2 \\ &= \sum_{\beta \in \{\alpha\}} (r_{\alpha,\beta})^2 (\alpha(\mu_\alpha))^2 \\ &= r(\mu_\alpha, \mu_\alpha)^2\end{aligned}$$

where r is rational and positive. Thus (μ_α, μ_α) is non-negative, and $\beta(\mu_\alpha) = (\mu_\alpha, \mu_\beta) = r_{\alpha,\beta}(\mu_\alpha, \mu_\alpha)$ is real. Q.E.D.

14. Definition. If Λ is a semi-simple Lie algebra and M a Cartan subalgebra, M_r is the real subspace of M spanned by the vectors μ_α, $\alpha \in \{\alpha\}$, i.e., the smallest subset of M closed under addition and under multiplication by real numbers which contains all the vectors μ_α, $\alpha \in \{\alpha\}$.

Note what is obvious: M_r is a real vector space.

15. Corollary. The dimension of the real vector space M_r is equal to the dimension of the complex vector space M.

Proof: Let m_r and m be these two dimensions. Then since by Corollary 10 M is spanned by vectors in M_r, it is clear that $m_r \geqslant m$. On the other hand, using Corollary 10, let $\mu_{\alpha_1}, \ldots, \mu_{\alpha_m}$ be a basis for M. If $\alpha \in \{\alpha\}$, we may then write $\mu_\alpha = c_1\mu_{\alpha_1} + \cdots + c_m\mu_{\alpha_m}$ with complex coefficients c_j. By Corollary 3, the matrix $(\mu_{\alpha_j}, \mu_{\alpha_i})$ is non-singular. Thus the system of equations

$$(\mu_\alpha, \mu_{\alpha_i}) = \sum_{j=1}^{m} c_j(\mu_{\alpha_j}, \mu_{\alpha_i}), \quad i = 1, \ldots, m$$

may be solved by Cramer's rule. Using Lemma 13, we see that all the coefficients c_j are real. Thus $m_r \leqslant m$. Q.E.D.

For the remainder of the present section, we let l denote the dimension of Λ, and m the dimension of M (which, as we have just seen, equals the dimension of M_r).

16. Lemma. The symmetric bilinear form (μ, ν) in the space M_r is real, and $(\mu, \mu) > 0$ if $\mu \in M_r$ is non-zero. Thus M_r taken with the inner product (μ, ν), is an m dimensional real Euclidean space.

Proof. Lemma 13 and Definition 14 show that (μ, ν) is real on M_r. We saw in the final paragraph of the proof of Lemma 4 that

$$(\mu, \mu) = \sum_{\alpha \in \{\alpha\}} (\alpha(\mu))^2 \dim (\Lambda_\alpha).$$

By Lemma 13, Lemma 4, and Corollary 8 this is a sum of non-negative terms, not all of which are zero. Q.E.D.

17. Definition. A set $\Pi \subseteq \{\alpha\}$ is called a *simple set of roots* if

(a) The set of linear functionals $\alpha \in \Pi$ (defined on M_r) are linearly independent.

(b) Every root $\beta \in \{\alpha\}$, $\beta \neq 0$, may be written as $\beta = \sum_{\alpha \in \Pi} k_\alpha \alpha$, where the coefficients k_α are unique, all integers, and are either all non-positive or all non-negative.

18. Lemma. There always exists a simple set of roots.

Proof. Let M_r' be the set of real-valued linear functionals β on the space M_r, or, equivalently, the set of complex valued linear functionals β on the space M which assume exclusively real values on the subset M_r of M. By Lemma 13, Corollary 3, and Corollary 8, every root belongs to M_r'. Introduce some basis $\beta_1, \ldots, \beta_m$ into the real vector space M_r', and using this basis, make M_r into an ordered vector space by writing $\gamma_1 > \gamma_2$ if the first non-zero coefficient of the vector $\gamma_1 - \gamma_2$ relative to the basis $\beta_1, \ldots, \beta_m$ is positive. In this way, M_r' is totally ordered, and we have $\gamma_1 + \gamma_2 > 0$ if $\gamma_1 > 0$ and $\gamma_2 > 0$, $-\gamma < 0$ if $\gamma > 0$. Call a vector $\gamma \in M_r'$ *positive* if $\gamma > 0$. Call $\mu_0 \in M_r$ positive if the $\gamma \in M_r'$ defined by $\gamma(\mu) = (\mu, \mu_0)$ is positive.

Let A_+ be the set of positive roots, A_- the set of negative roots, so that $A_- = -A_+$, and $\{\alpha\} = A_+ \cup A_- \cup \{0\}$. Call a positive root *decomposable* if $\beta = \alpha_1 + \alpha_2$, with $\alpha_1, \alpha_2 > 0$, $\alpha_1, \alpha_2 \in \{\alpha\}$. Let Π be the collection of all indecomposable positive roots.

Since $\beta = \alpha_1 + \alpha_2$ with $\alpha_1 > 0$ and $\alpha_2 > 0$ implies $\beta > \alpha_1$ and $\beta > \alpha_2$, it is plain that each positive root may be written as $\beta = \sum_{\alpha \in \Pi} k_\alpha \alpha$, where each k_α is a non-negative integer. Thus, using Corollary 2, it follows that each negative root may be written as $\beta = \sum_{\alpha \in \Pi} k_\alpha \alpha$, where each k_α is a non-positive integer. This makes it clear that both (a) and (b) of Definition 14 will follow by Corollary 8 if only we can show that the vectors μ_α, $\alpha \in \Pi$, are linearly independent.

This we prove as follows. If $\alpha, \beta \in \Pi$, $\alpha \neq \beta$, and $\alpha - \beta$ is a root, then either $\alpha - \beta$ or $\beta - \alpha$ is a positive root. Suppose to be specific that $\alpha - \beta = \gamma$ is a positive root. Then $\alpha = \beta + \gamma$ is decomposable, a contradiction. This shows that if $\alpha, \beta \in \Pi$ and $\alpha \neq \beta$, $\alpha - \beta$ is not a root. By Corollary 12 and Corollary 8, $(\mu_\alpha, \mu_\beta) \leqslant 0$. Now, suppose that we had a linear relation $\mu_\alpha = \sum_{\substack{\beta \in \Pi \\ \beta \neq \alpha}} c_\beta \mu_\beta$ with real coefficients c_β. Let Π_+ be the subset of Π for which $c_\beta > 0$, and Π_- the subset of Π for which $c_\beta < 0$, so that

$$\mu_\alpha = \sum_{\beta \in \Pi_+} d_\beta \mu_\beta - \sum_{\beta \in \Pi_-} d_\beta \mu_\beta = \mu^+ - \mu^-,$$

with real positive coefficients d. Since $\mu_\beta > 0$ for all $\beta \in \Pi$, we cannot have $\mu^+ = 0$. But therefore, by Lemma 16 and using $(\mu_\beta, \mu_\gamma) \leqslant 0$ for $\beta \neq \gamma$, we have

$$(\mu_\alpha, \mu^+) = (\mu^+, \mu^+) - \sum_{\substack{\beta \in \Pi_- \\ \gamma \in \Pi_+}} d_\beta d_\gamma (\mu_\beta, \mu_\gamma) > 0.$$

On the other hand, for the same reason

$$(\mu_\alpha, \mu^+) = \sum_{\gamma \in \Pi_+} d_\gamma (\mu_\alpha, \mu_\gamma) \leqslant 0,$$

a contradiction. This shows that the vectors μ_α, $\alpha \in \Pi$, are linearly independent, and completes our proof. Q.E.D.

19. Corollary. Let Π be a simple set of roots. Then

(i) The vectors μ_α, $\alpha \in \Pi$, form a basis for M_r;

(ii) Every root $\beta \neq 0$ may be built up from roots $\alpha \in \Pi$ and their negatives by an iterated sequence of additions $\beta_1 + \beta_2$, both terms β_1 and β_2 always being non-zero roots, and $\beta_1, \beta_2 \in \Pi \cup -\Pi$;

(iii) The smallest Lie subalgebra of Λ containing all the spaces Λ_α and $\Lambda_{-\alpha}$, $\alpha \in \Pi$, is Λ itself.

(iv) Every element of Λ may be written as a linear combination of elements of the form $[[\ldots [\lambda_{\alpha_1}, \lambda_{\alpha_2}], \ldots] \lambda_{\alpha_n}]$, where $\alpha_j \in \pm\Pi$, $j = 1, \ldots, n$.

Proof. By Definition 17, the roots $\alpha \in \Pi$ form a basis for the space M_r' of the preceding proof. Suppose that we use *this* basis to introduce a partial ordering into M_r', as in the preceding proof. Then, as that proof shows, the set of all positive roots which are indecomposable in the sense of the preceding proof forms a simple set of roots, and hence a linearly independent set of roots. But it is plain from Definition 17 that each $\alpha \in \Pi$ is indecomposable. Since a basis cannot be enlarged and still remain linearly independent, it follows that Π is identical with the set of all indecomposable roots. But then it follows immediately that every root $\beta > 0$ (resp $\beta < 0$) can be built up from the roots $\alpha \in \Pi$ (resp $\alpha \in -\Pi$) by a series of additions $\beta_1 + \beta_2$. It follows, using Lemma 11, that if $\lambda_\beta \in \Lambda_\beta$ where $\beta > 0$, λ_β belongs to the smallest Lie subalgebra $\tilde{\Lambda}$ of Λ containing all the vectors λ_α, $\alpha \in \Pi$. Similarly, using Corollary 2, $\Lambda_\beta \subseteq \tilde{\Lambda}$ if $\beta \in \{\alpha\}$ and $\beta < 0$. Thus $\tilde{\Lambda} \supseteq \Lambda_\beta$ if $\beta \neq 0$ and $\beta \in \{\alpha\}$. Hence $\tilde{\Lambda}$ contains all the vectors μ_β, so, using Corollary 10, $\tilde{\Lambda} \supseteq \Lambda_0$, and thus $\tilde{\Lambda} \supseteq \sum_{\beta \in \{\alpha\}} \oplus \Lambda_\beta = \Lambda$. This proves (iii).

By the Jacobi identity, if $\lambda_1, \ldots, \lambda_p, \lambda'_1, \ldots, \lambda'_q \in \Lambda$, we have

$$\begin{aligned}
&[[[\cdots[\lambda_1, \lambda_2], \ldots], \lambda_p], [[\cdots[\lambda'_1, \lambda'_2], \ldots], \lambda'_q]] \\
&\quad = -[[[[\cdots[\lambda_1, \lambda_2], \ldots], \lambda_p], \lambda'_q], [[\cdots[\lambda'_1, \lambda'_2], \ldots], \lambda'_{q-1}]] \\
&\quad\quad + [[[[[\cdots[\lambda_1, \lambda_2]\ldots], \lambda_p], [[\cdots[\lambda'_1, \lambda'_2], \ldots], \lambda'_{q-1}]], \lambda'_q].
\end{aligned}$$

It therefore follows inductively using what we have already proved, that if $\beta > 0$ and $\lambda_\beta \in \Lambda_\beta$, then λ_β may be written as a linear combination of elements $[[\ldots[\lambda_{\alpha_1}, \lambda_{\alpha_2}], \ldots]\, \lambda_{\alpha_n}]$, with $\alpha_1, \ldots, \alpha_n \in \Pi$. This proves (iv). Moreover, by Lemma 11, it follows that there exists a sequence $\alpha_1, \ldots, \alpha_n \in \Pi$ such that $\beta = \alpha_1 + \cdots + \alpha_n$ and such that $\alpha_1 + \cdots + \alpha_k$ is a positive root for each $k > 0$. Statement (ii) now follows immediately by Corollary 2.

Statement (i) follows by Corollary 10, Lemma 8, Corollary 3, and Definition 17. Q.E.D.

In formulating the following basic result on isomorphism we let Λ and Λ' be two semi-simple complex Lie algebras, and M and M' Cartan subalgebras of Λ and Λ' respectively. Let M_r and M'_r be the real subspaces of M and M' respectively described by Definition 14. We let (λ, μ) be the natural bilinear form of Λ, and (λ', μ') that of Λ'. For each root α of M in Λ, we let Λ_α be the corresponding root space, and let μ_α be the vector described by Corollary 8; similarly, for each root α of M' in Λ', we let Λ'_α be the corresponding root space, and let μ'_α be the vector described by Corollary 8.

20. Lemma. Suppose, assuming the notations explained just above, that Π and Π' are simple sets of roots of Λ and Λ' respectively. Let ϕ be a one-to-one linear mapping of the real vector space M_r into the real vector space M'_r, such that

(i) For each $\alpha \in \Pi$, $\phi(\mu_\alpha) = \mu'_{\alpha'}$, where $\alpha' \in \Pi'$.
(ii) The transformation $\alpha \to \alpha'$ defined by (i) maps Π onto Π'.
(iii) We have $(\phi(\mu_\alpha), \phi(\mu_\beta)) = (\mu_\alpha, \mu_\beta)$ for $\alpha, \beta \in \Pi$.

Then Λ and Λ' are isomorphic; indeed, ϕ may be extended to an isomorphism of Λ onto Λ'.

Proof. For each $\alpha \in \Pi \cup (-\Pi)$ we choose a vector $\lambda_\alpha \in \Lambda_\alpha$, making sure that $(\lambda_\alpha, \lambda_{-\alpha}) = 1$, so that $[\lambda_\alpha, \lambda_{-\alpha}] = \mu_\alpha$ by Corollary 8. Similarly, for each $\alpha' \in \Pi' \cup (-\Pi')$, we let $\lambda'_{\alpha'} \in \Lambda'_{\alpha'}$, and suppose $(\lambda'_{\alpha'}, \lambda'_{-\alpha'}) = 1$, so that $[\lambda'_{\alpha'}, \lambda'_{-\alpha'}] = \mu'_{\alpha'}$. We shall first prove

(a) If $\alpha_0, \dots, \alpha_p$ is a sequence of elements of $\Pi \cup (-\Pi)$ such that $\alpha_0 + \cdots + \alpha_p = 0$, then we have

$$\phi([[\dots[[\lambda_{\alpha_0}, \lambda_{\alpha_1}], \lambda_{\alpha_2}]\dots], \lambda_{\alpha_p}]) = [[\dots[[\lambda'_{\alpha'_0}, \lambda'_{\alpha'_1}], \lambda'_{\alpha'_2}], \dots]\, \lambda'_{\alpha'_p}].$$

To prove (a), define the rank of the sequence $\alpha_0, \dots, \alpha_p$ as the pair $[p, k]$, where

$$k = \sum_{\alpha_i \in \Pi} i \quad \text{if} \quad \alpha_0 \in (-\Pi)$$

$$k = \sum_{\alpha_i \in (-\Pi)} i \quad \text{if} \quad \alpha_0 \in \Pi.$$

If $p = i$, (a) reduced to $\phi([\lambda_\alpha, \lambda_{-\alpha}]) = [\lambda'_{\alpha'}, \lambda'_{-\alpha'}]$, i.e. to $\phi(\mu_\alpha) = \mu'_{\alpha'}$, which is simply (i). We shall suppose that (a) is known for each sequence $\hat{a}_0 \cdots \hat{a}_q \in \Pi \cup (-\Pi)$ and such that $\hat{a}_0 + \cdots + \hat{a}_q = 0$ of rank $[q, k']$ such that either $q < p$ or $q = p$ and $k' < k$, and show that this implies (a) for the sequence $\alpha_0, \dots, \alpha_p$. This double induction will evidently prove the general validity of (a). We may assume, of course, that $p > 1$, and for the sake of definiteness, that $\alpha_0 \in \Pi$. By Corollary 19 it follows that the sequence $\alpha_0, \dots, \alpha_p$ contains as many elements of Π as of $(-\Pi)$. Thus $\alpha_0, \dots, \alpha_p$ must contain a pair α_j, α_{j+1} of successive elements such that $\alpha_j \in (\Pi)$, $\alpha_{j+1} \in -\Pi$. Unless $\alpha_j = -\alpha_{j+1}$, λ_{α_j} and $\lambda_{\alpha_{j+1}}$ commute by Definition 17 and Lemma 2.5 so that

$$[\dots[[\lambda_{\alpha_0}, \lambda_{\alpha_1}], \lambda_{\alpha_2}]\dots], \lambda_{\alpha_p}] = 0 \quad \text{if} \quad j = 0, \quad \text{and}$$

$$\begin{aligned} &[\dots[[\lambda_{\alpha_0}, \lambda_{\alpha_1}], \lambda_{\alpha_2}]\dots], \lambda_{\alpha_j}], \lambda_{\alpha_{j+1}}]\dots], \lambda_{\alpha_p}] \\ &\quad = [\dots[[\lambda_{\alpha_0}, \lambda_{\alpha_1}], \lambda_{\alpha_2}]\dots], \lambda_{\alpha_{j+1}}], \lambda_{\alpha_j}]\dots], \lambda_{\alpha_p}] \end{aligned}$$

The corresponding identity holds for the vectors $\lambda'_{\alpha'_0}, \dots, \lambda'_{\alpha'_p}$, so that in case $\alpha_j \neq -\alpha_{j+1}$ (a) follows by inductive hypothesis. If $\alpha_j = -\alpha_{j+1}$, we have $[\lambda_{\alpha_j}, \lambda_{\alpha_{j+1}}] = \mu_{\alpha_j}$, so that using the Jacobi identity we find

$$\begin{aligned} &[\dots[[\lambda_{\alpha_0}, \lambda_{\alpha_1}], \lambda_{\alpha_2}]\dots, \lambda_{\alpha_j}]\, \lambda_{\alpha_{j+1}}]\dots], \lambda_{\alpha_p}] \\ &\quad = [\dots[[\lambda_{\alpha_0}, \lambda_{\alpha_1}], \lambda_{\alpha_2}]\dots], \lambda_{\alpha_{j+1}}], \lambda_{\alpha_j}]\dots], \lambda_{\alpha_p}] \\ &\qquad + (\alpha_0 + \cdots + \alpha_{j-1})(\mu_{\alpha_j})\, [\dots[[\lambda_{\alpha_0}, \lambda_{\alpha_1}], \lambda_{\alpha_2}]\dots, \lambda_{\alpha_{j-1}}], \lambda_{\alpha_{j+2}}]\dots, \lambda_{\alpha_p}], \end{aligned}$$

and a corresponding identity for the vectors $\lambda'_{\alpha'_0}, \dots, \lambda'_{\alpha'_p}$. Since by hypothesis and by Corollary 8 we have $\alpha_k(\mu_{\alpha_j}) = (\mu_{\alpha_j}, \mu_{\alpha_k}) = (\mu'_{\alpha'_j}, \mu'_{\alpha'_k})$

$= \alpha'_k(\mu'_{\alpha'_j})$, (a) also follows in this case by inductive hypothesis. Thus the proof of (a) is complete.

Next we prove the following statement:

(b) If $\alpha'_0, \ldots, \alpha'_p$ and $\beta'_0, \ldots, \beta'_q$ are elements of $\Pi \cup (-\Pi)$ such that $\alpha'_0 + \cdots + \alpha'_p + \beta'_0 + \cdots + \beta'_q = 0$, and if we put

$$[\alpha'_0, \ldots, \alpha'_p] = [\ldots [[\lambda'_{\alpha'_0}, \lambda'_{\alpha'_1}]\, \lambda'_{\alpha'_2}] \ldots]\, \lambda'_{\alpha'_p}]$$

and define $[\beta'_0, \ldots, \beta'_q]$, $[\alpha_0, \ldots, \alpha_p]$, $[\beta_0, \ldots, \beta_q]$ similarly, we have

$$[[\alpha'_0, \ldots, \alpha'_p], [\beta'_0, \ldots, \beta'_q]] = \phi([\alpha_0, \ldots, \alpha_p], [\beta_0, \ldots, \beta_q]).$$

It is easiest to prove (b) by proving the following slightly more general assertion:

(b′) If $\alpha_0, \ldots, \alpha_p, \beta_0, \ldots, \beta_q$, and $\gamma_0, \ldots, \gamma_r$ are elements of $\Pi \cup (-\Pi)$ such that $\alpha_0 + \cdots + \alpha_p + \beta_0 + \cdots + \beta_q + \gamma_0 + \cdots + \gamma_r = 0$, then

$$\begin{aligned} &\phi([[\ldots [[[\alpha_0, \ldots, \alpha_p], [\beta_0, \ldots, \beta_q]]\, \lambda_{\gamma_0}] \ldots]\, \lambda_{\gamma_r}]) \\ &\qquad = [[\ldots [[[\alpha'_0, \ldots, \alpha'_p], [\beta'_0, \ldots, \beta'_q]]\, \lambda'_{\gamma'_0}, \ldots]\, \lambda'_{\gamma'_r}]. \end{aligned}$$

To prove (b'), we have only to note that by the Jacobi identity

$$\begin{aligned} &[[\ldots [[[\alpha_0, \ldots, \alpha_p], [\beta_0, \ldots, \beta_q]]\, \lambda_{\gamma_0}] \ldots]\, \lambda_{\gamma_r}] \\ &\quad = -[[\ldots [[[\alpha_1, \ldots, \alpha_{p-1}], [\beta_0, \ldots, \beta_q, \alpha_p]]\, \lambda_{\gamma_0}] \ldots]\, \lambda_{\gamma_r}] \qquad (1) \\ &\qquad - [[\ldots [[[[\alpha_0, \ldots, \alpha_{p-1}], [\beta_0, \ldots, \beta_q]]\, \lambda_{\alpha_p}]\, \lambda_{\gamma_0}] \ldots]\, \lambda_{\gamma_q}], \end{aligned}$$

a similar identity holding for the vectors $\lambda'_{\alpha'}$. Thus (b′) follows by induction on p from its special case $p = 1$, which is simply (a).

From (b), it is easy to prove the following statement (c), which is formulated in terms of the notations introduced above:

(c) If $\alpha_0, \ldots, \alpha_p$ and $\beta_0, \ldots, \beta_q$ are arbitrary sequences of elements of $\Pi \cup (-\Pi)$ such that $\alpha_0 + \cdots + \alpha_p = \beta_0 + \cdots + \beta_q \neq 0$, then the ratio of the parallel vectors

$$[\alpha_0, \ldots, \alpha_p] \quad \text{and} \quad [\beta_0, \ldots, \beta_q]$$

is the same as that of the vectors

$$[\alpha'_0, \ldots, \alpha'_p] \quad \text{and} \quad [\beta'_0, \ldots, \beta'_q]$$

If $\alpha_0 + \cdots + \alpha_p = 0$, statement (c) is merely (a). Let $\alpha_0 + \cdots + \alpha_p = \gamma \neq 0$. By Corollary 19 (iii) and Lemma 11, there exists a sequence

$\gamma_1, \dots, \gamma_r$ of elements of $\Pi \cup -\Pi$ such that $\gamma_1 + \cdots + \gamma_r = \gamma$ and such that $[\gamma_1, \dots, \gamma_r] \neq 0$. It then follows by (b) that the ratio of the parallel vectors

$$[[\alpha_0, \dots, \alpha_p], [\gamma_1, \dots, \gamma_r]] \quad \text{and} \quad [[\beta_1, \dots, \beta_q], [\gamma_1, \dots, \gamma_r]]$$

is the same as that of the vectors

$$[[\alpha'_0, \dots, \alpha'_p], [\gamma'_1, \dots, \gamma'_r]] \quad \text{and} \quad [[\beta'_1, \dots, \beta'_q], [\gamma'_1, \dots, \gamma'_r]].$$

Statement (c) now follows from Corollary 8.

It is now easy to complete the proof of our Lemma. We extend ϕ to a mapping Φ defined on all of Λ be requiring that $\Phi([\alpha_1, \dots, \alpha_p]) = [\alpha'_1, \dots, \alpha'_p]$ for all $\alpha_1, \dots, \alpha_p \in \Pi \cup -\Pi$. By Corollary 19 and (c), Φ is a linear transformation mapping Λ onto Λ'. The symmetry of the relationship between Λ and Λ' makes it plain that Φ has an inverse Ψ defined by $\Psi([\alpha'_1, \dots, \alpha'_p]) = [\alpha_1, \dots, \alpha_p]$. Thus Φ is one-to-one. Using the identity 1, it is easy to prove by induction on q that

$$\Phi([\alpha_1, \dots, \alpha_p], [\beta_1, \dots, \beta_q]) = [\Phi([\alpha_1, \dots, \alpha_p]), \Phi([\beta_1, \dots, \beta_q])].$$

Thus Φ is an isomorphism of Λ with Λ'. Q.E.D.

The following theorem recasts Lemma 20 in a form which will facilitate our subsequent calculations.

21. Theorem. Suppose, assuming the notations explained in the paragraph immediately preceding Lemma 20, that Π and Π' are simple sets of roots of Λ and Λ' respectively. Let $\alpha \to \alpha'$ be a one-to-one transformation of Π onto Π'. Suppose that for each non-negative integer n, and each $\alpha, \beta \in \Pi$, $\alpha + n\beta$ is a root of M in Λ if and only if $\alpha' + n\beta'$ is a root of M' in Λ'. Then Λ and Λ' are isomorphic.

Proof. We shall show that the hypotheses of the present theorem imply the hypotheses of Lemma 20. By Definition 17, each root $\beta \in \{\alpha\}$, $\beta \neq 0$, can be written as $\beta = \sum_{\alpha \in \Pi} k_\alpha \alpha$, where all the k_α are integers, and all either non-positive or non-negative. Similarly, each non-zero root $\tilde{\beta}$ of M' in Λ' can be written as $\tilde{\beta} = \sum_{\alpha' \in \Pi'} k'_{\alpha'} \alpha'$, where all the $k'_{\alpha'}$ are integers, and all either non-positive or non-negative. We call β (resp. $\tilde{\beta}$) positive if all the k_α (resp. $k'_{\alpha'}$) are non-negative, negative otherwise. We call $\sum_{\alpha \in \Pi} |k_\alpha|$ (resp. $\sum_{\alpha' \in \Pi'} |k'_{\alpha'}|$) the rank of β (resp. $\tilde{\beta}$).

We first show

(a) If $\beta = \sum_{\alpha \in \Pi} k_\alpha \alpha$ is a root, then $\sum_{\alpha' \in \Pi'} k_\alpha \alpha'$ is a root.

This we prove by induction on the rank of β. If this rank is 1, we merely have one of our hypotheses. Suppose that the rank of β is $k > 1$, and that (a) is known for all roots whose rank is less than k. We suppose for the sake of definiteness that β is positive. By Corollary 19, there exists an $\alpha_0 \in \Pi$ such that $\beta_0 = \beta - \alpha_0$ is a root. Put $n_\alpha = k_\alpha$ if $\alpha \neq \alpha_0$, $n_{\alpha_0} = k_{\alpha_0} - 1$, $\beta' = \sum_{\alpha' \in \Pi'} k_\alpha \alpha'$, $\beta'_0 = \sum_{\alpha' \in \Pi'} n_\alpha \alpha'$. Then, by inductive hypothesis, $\beta_0 + n\alpha$ and $(\sum_{\alpha' \in \Pi'} n_\alpha \alpha') + n\alpha'$ are either both roots or both not roots if $n \leqslant 0$. Let n_- be the non-positive n, largest in absolute value, for which $\beta_0 + n\alpha$ is a root. It follows by Corollary 8 and Corollary 12 that β is a root if and only if

$$-\sum_{\alpha \in \Pi} \frac{(\mu_\alpha, \mu_{\alpha_0})}{(\mu_{\alpha_0}, \mu_{\alpha_0})} = -\frac{(\mu_{\beta_0}, \mu_{\alpha_0})}{(\mu_{\alpha_0}, \mu_{\alpha_0})} > \frac{1}{2} n_-,$$

and β' is a root if and only if

$$-\sum_{\alpha' \in \Pi'} n_\alpha \frac{(\mu_{\alpha'}, \mu_{\alpha'_0})}{(\mu_{\alpha'_0}, \mu_{\alpha'_0})} = -\frac{(\mu'_{\beta'_0}, \mu'_{\alpha'_0})}{(\mu'_{\alpha'_0}, \mu'_{\alpha'_0})} > \frac{1}{2} n_-.$$

But, by hypothesis, and using Corollary 8 and Corollary 12 once more,

$$\frac{(\mu_\alpha, \mu_{\alpha_0})}{(\mu_{\alpha_0}, \mu_{\alpha_0})} = \frac{(\mu'_{\alpha'}, \mu'_{\alpha'_0})}{(\mu'_{\alpha'_0}, \mu'_{\alpha'_0})}, \alpha \in \Pi.$$

Thus β is a root if and only if β' is a root, proving (a).

Put $\beta' = \sum_{\alpha' \in \Pi'} k_\alpha \alpha'$ for each $\beta = \sum_{\alpha \in \Pi} k_\alpha \alpha \in \{\alpha\}$. Then from (a) and Corollary 12, we deduce that

(b) $$\frac{(\mu_\alpha, \mu_\beta)}{(\mu_\beta, \mu_\beta)} = \frac{(\mu'_{\alpha'}, \mu'_{\beta'})}{(\mu'_{\beta'}, \mu'_{\beta'})}, \quad \alpha, \beta \in \{\alpha\}, \ \alpha, \beta \neq 0.$$

We saw in the final paragraph of the proof of Lemma 4 that

$$(\mu, \mu) = \sum_{\alpha \in \{\alpha\}} (\alpha(\mu))^2 \dim(\Lambda_\alpha), \qquad \mu \in M$$

so that by Corollary 8 and Lemma 7

$$(\mu, \mu) = \sum_{\alpha \in \{\alpha\}} (\mu_\alpha, \mu)^2, \qquad \mu \in M$$

Thus

$$(\mu_\beta, \mu_\beta)^{-1} = \sum_{\alpha \in \{\alpha\}} \frac{(\mu_\alpha, \mu_\beta)^2}{(\mu_\beta, \mu_\beta)^2}$$

and similarly

$$(\mu'_{\beta'}, \mu'_{\beta'})^{-1} = \sum_{\alpha' \in \{\alpha'\}} \frac{(\mu'_{\alpha'}, \mu'_{\beta'})^2}{(\mu'_{\beta'}, \mu'_{\beta'})^2}$$

But then, by (b), we have

$$(\mu'_{\beta'}, \mu'_{\beta'}) = (\mu_{\beta}, \mu_{\beta}), \ \beta \in \{\alpha\}$$

Using (b) once more, the hypotheses of Lemma 20 are verified.

Now define the linear transformation ϕ on M_r by $\phi(\mu_\alpha) = \mu'_{\alpha'}$, $\alpha \in \Pi$ (c.f. Definition 17). Then the hypotheses of Lemma 20 are satisfied, and our theorem is proved.

The following corollary gives useful auxiliary information about the isomorphism whose existence is asserted by Theorem 21.

22. Corollary. Suppose, under the hypotheses of the preceding theorem, that for each $\alpha \in \Pi$ we choose vectors λ_α, $\lambda_{-\alpha}$, $\lambda'_{\alpha'}$, $\lambda'_{-\alpha'}$, such that $\lambda_{\pm\alpha} \in \Lambda_{\pm\alpha}$, $\lambda'_{\pm\alpha'} \in \Lambda'_{\pm\alpha'}$ in such a way that $(\lambda_\alpha, \lambda_{-\alpha}) = (\lambda'_{\alpha'}, \lambda'_{-\alpha'}) = 1$. Then the formula

$$\Phi[[\ldots [[\lambda_{\alpha_1}, \lambda_{\alpha_2}], \lambda_{\alpha_3}] \ldots], \lambda_{\alpha_n}] = [[\ldots [[\lambda'_{\alpha'_1}, \lambda'_{\alpha'_2}], \lambda'_{\alpha'_3}] \ldots], \lambda'_{\alpha'_n}]$$

$\alpha_1, \ldots, \alpha_n \in \Pi \cup -\Pi$, defines an isomorphism of Λ onto Λ'.

Proof. Examination of the first and of the final paragraph of the proof of Lemma 20 will show that the isomorphism between Λ and Λ' constructed in the course of that proof has the properties described in the present corollary. Q.E.D.

Theorem 21 shows that the structure of Λ is entirely determined by the geometric configuration of the m vectors constituting the set $\{\mu_\alpha\}$, $\alpha \in \Pi$, in the m-dimensional Euclidean space M_r. Analysis of this geometric configuration now leads rapidly to a complete understanding of the structure of Λ.

Throughout the remainder of the present section, we let Π be a simple set of roots.

23. Lemma. If $\alpha, \beta \in \Pi$ and $\alpha \neq \beta$, then

$$[*]\ n_{\alpha,\beta} = 4 \frac{(\mu_\alpha, \mu_\beta)^2}{(\mu_\alpha, \mu_\alpha)(\mu_\beta, \mu_\beta)} = 0, 1, 2, \text{or } 3.$$

Moreover, if $n_{\alpha,\beta} \neq 0$, then one of

$$2\,\frac{(\mu_\alpha, \mu_\beta)}{(\mu_\alpha, \mu_\alpha)} \text{ and } 2\,\frac{(\mu_\alpha, \mu_\beta)}{(\mu_\beta, \mu_\beta)}$$

is -1, and the other is $-n_{\alpha,\beta}$.

Proof. By Corollaries 12 and 8, $n_{\alpha,\beta}$ is an integer. By the Schwarz inequality, this integer is at most 4; and, if it is 4, we must have $\mu_\alpha = -\mu_\beta$. But then, by Corollary 8, $\alpha = -\beta$, is impossible since $\alpha, \beta \in \Pi$ by Definition 17(b). Thus $n_{\alpha,\beta} \leqslant 3$, and all our assertions are evident (cf. Corollary 12). Q.E.D.

For the rest of present section, if $\alpha, \beta \in \Pi$, we shall use $n_{\alpha,\beta}$ for the integer defined by formula [*].

24. Definition. If $\alpha, \beta \in \Pi$ and $\alpha \neq \beta$, we call α and β *neighbors* if $n_{\alpha,\beta} \neq 0$, *and simple neighbors, double neighbors, or triple neighbors* according as $n_{\alpha,\beta}$ is 1, 2, or 3. If α and β are double (or triple) neighbors, we say that the *doubling (or tripling) is on* β if $\alpha + 2\beta$ (or $\alpha + 3\beta$) is a root of μ in Λ.

We now begin to analyze details of the geometric configuration of the vectors μ_α, $\alpha \in \Pi$.

25. Lemma. There exists no sequence $\alpha_1, \alpha_2, \ldots, \alpha_n$ $n \geqslant 3$ of distinct elements of Π such that α_i is a neighbor of α_{i+1} and α_n a neighbor of α_1.

Proof. Write $u_\alpha = \mu_\alpha/|\mu_\alpha|$ if $\alpha \in \Pi$, where $|\mu_\alpha|$ is the length of μ_α. Then since the vectors μ_α are linearly independent (cf. Definition 17) we have

$$0 < \Big|\sum_i u_{\alpha_i}\Big|^2 = \sum_{i,j} (u_{\alpha_i}, u_{\alpha_j}) = n - \sum_{\substack{i<j \\ j=2,\cdots,n}} \sqrt{n_{\alpha_i,\alpha_j}} \leqslant 0$$

Q.E.D.

26. Lemma. Let the neighbors of an $\alpha \in \Pi$ be n in number, where we agree to count a double neighbor twice and a triple neighbor 3 times. Then $n \leqslant 3$.

Proof. Let the distinct neighbors of α be $\alpha_1, \ldots, \alpha_p$. As in the proof of the preceding lemma, write $u_\alpha = \mu_\alpha/|\mu_\alpha|$. By the preceding lemma, the vectors $u_{\alpha_1}, \ldots, u_{\alpha_p}$ are orthogonal unit vectors. By Definition 17, the vectors $u_\alpha, u_{\alpha_1}, \ldots, u_{\alpha_p}$ are linearly independent. Thus, by Pythagoras' theorem, $1 > \sum_{i=1}^{p} (u_\alpha, u_{\alpha_i})^2 = \frac{1}{4}\sum_{i=1}^{p} n_{\alpha,\alpha_i}$, so that $\sum_{i=1}^{p} n_{\alpha,\alpha_i} \leqslant 3$. Q.E.D.

From this point to the end of the present section, we wish to restrict our attention to simple complex Lie algebras. The following Lemma gives an easily applied criterion for a semi-simple algebra to be simple.

27. Lemma. The *semi-simple* Lie algebra Λ is simple if and only if any two elements α and β of Π can be connected by a sequence

$$\alpha = \alpha_1, \alpha_2, \dots, \alpha_{n-1}, \quad \alpha_n = \beta$$

of elements of Π, each of which is a neighbor of the next.

Proof. First suppose that Π does not have this property, so that Π may be divided into two disjoint subset $\Pi = \Pi_1 \cup \Pi_2$, in such a way that no $\alpha \in \Pi_1$ is the neighbor of any $\beta \in \Pi_2$. Then, by Definition 24, Definition 17, and Lemma 11, we have $[\Lambda_{\pm\alpha}, \Lambda_{\pm\beta}] = 0$ if $\alpha \in \Pi_1$ and $\beta \in \Pi_2$. It follows readily by induction that if Λ_1 (resp. Λ_2) is the Lie subalgebra of Λ generated by all the subspaces $\Lambda_{\pm\alpha}$, $\alpha \in \Pi_1$ (resp. $\alpha \in \Pi_2$), we have $[\Lambda_1, \Lambda_2] = 0$. Thus $\Lambda_1 + \Lambda_2$ is a subalgebra of Λ, and, by Corollary 19 (iii) $\Lambda_1 + \Lambda_2 = \Lambda$. Then plainly Λ_1 and Λ_2 are ideals in Λ. Were Λ simple, we would have $\Lambda_1 = \Lambda_2 = \Lambda$, and therefore $[\Lambda, \Lambda] = 0$, an impossibility. This shows that if Π is not connected in the sense of the present lemma, then Λ is not simple.

Conversely, suppose that Λ is not simple, so that by Theorem 2.17 we may write $\Lambda = \Lambda^{(1)} \oplus \Lambda^{(2)}$, where $\Lambda^{(1)}$ and $\Lambda^{(2)}$ are ideals of Λ. Both $\Lambda^{(1)}$ and $\Lambda^{(2)}$ are invariant under the operators Ad (μ), where $\mu \in M$; hence, by Theorem 1.10, $\Lambda^{(1)} = \sum_{\alpha \in A^{(1)}} \oplus \Lambda_\alpha^{(1)}$, and $\Lambda^{(2)} = \sum_{\alpha \in A^{(2)}} \oplus \Lambda_\alpha^{(2)}$, where $A^{(1)}$ and $A^{(2)}$ are respectively the sets of weights of the representations $\mu \to \text{Ad}(\mu)|\Lambda^{(1)}$ and $\mu \to \text{Ad}(\mu)|\Lambda^{(2)}$ of M. Thus the decomposition of Λ into root spaces of M must be given by $\Lambda = \sum_{\alpha \in A^{(1)}} \oplus \Lambda_\alpha^{(1)} \oplus \sum_{\alpha \in A^{(2)}} \oplus \Lambda_\alpha^{(2)}$. Since the root space Λ_α corresponding to non-zero root α of M in Λ is one-dimensional, we must therefore have either $\Lambda_\alpha \subseteq \Lambda^{(1)}$ or $\Lambda_\alpha \subseteq \Lambda^{(2)}$ for all $0 \neq \alpha \in \{\alpha\}$, and in particular, for all $\alpha \in \Pi$. Since $[\Lambda_{-\alpha}, \Lambda_\alpha] \neq 0$, we cannot have $\Lambda_\alpha \subseteq \Lambda^{(1)}$ and $\Lambda_{-\alpha} \subseteq \Lambda^{(2)}$; thus if $\Lambda_\alpha \subseteq \Lambda^{(1)}$, $\Lambda_{-\alpha} \subseteq \Lambda^{(1)}$. If we let Π_1 (resp. Π_2) denote the set of $\alpha \in \Pi$ such that $\Lambda_\alpha \subseteq \Lambda^{(1)}$ (resp. $\Lambda^{(2)}$), it follows by Corollary 19 (iii) that neither Π_1, nor Π_2 is null. The set Π is plainly the disjoint union of Π_1 and Π_2; since $[\Lambda_{\pm\alpha}, \Lambda_{\pm\beta}] = 0$ if $\alpha \in \Pi_1$, and $\beta \in \Pi_2$, no α in Π_1 is the neighbor of any $\beta \in \Pi_2$. This completes the proof of the converse part of the present lemma. Q.E.D.

The following corollary disposes completely of the case in which any $\alpha \in \Pi$ has a triple neighbor.

28. Corollary. Let Λ be simple. If Π contains any element α with a triple neighbor β, it contains no elements other than α and β.

Proof. This follows immediately from Lemma 26 and Lemma 27. Q.E.D.

For the rest of the present section, we suppose that no $\alpha \in \Pi$ has a triple neighbor.

29. Lemma. Let Λ be simple. Let $\alpha \in \Pi$ have the double neighbor β. Then no element $\gamma \in \Pi$ other than α or β has either a double neighbor, or more than two single neighbors.

Proof. First suppose that there exists a $\gamma \in \Pi, \gamma \neq \alpha, \gamma \neq \beta$ with a double neighbor. Then, using Lemmas 25, 26 and 27, and interchanging α with β if necessary, we may find a sequence $\alpha_1 = \alpha, \alpha_2 = \beta, \alpha_3, \ldots, \alpha_{n-1}, \alpha_n$ of distinct elements of Π, such that α_1 and α_2 are double neighbors, α_{n-1} and α_n are double neighbors, α_i is a simple neighbor of α_{i-1} and α_{i+1} if $2 < i < n-1$, and no α_i has any neighbors α_j in the sequence but α_{i+1} and α_{i-1}. Put $u_1 = \mu_{\alpha_1}/|\mu_{\alpha_1}|$, $u_2 = (\mu_{\alpha_2} + \cdots + \mu_{\alpha_{n-1}})/|\mu_{\alpha_2} + \cdots + \mu_{\alpha_{n-1}}|$, and $u_3 = \mu_{\alpha_n}/|\mu_{\alpha_n}|$. This gives three unit vectors, linearly independent by Definition 17. By Lemma 23, u_1 and u_3 are orthogonal. We have $|\mu_{\alpha_2}|^2 = \cdots = |\mu_{\alpha_{n-1}}|^2$ and $(\mu_{\alpha_j}, \mu_{\alpha_{j+1}}) = -\frac{1}{2}|\mu_{\alpha_2}|^2$ for $2 < j < n-1$ by Lemma 23. Thus

$$|\mu_{\alpha_2} + \cdots + \mu_{\alpha_{n-1}}|^2 = |\mu_{\alpha_2}|^2 = |\mu_{x_{n-1}}|^2$$

and therefore $(u_2, u_3) = (u_1, u_2) = -\frac{1}{2}\sqrt{2}$ by Lemma 23. But then $1 = |u_2|^2 = (u_2, u_1)^2 + (u_2, u_3)^2$, contradicting the linear independence of u_1, u_2, u_3.

Next suppose that there exists a $\gamma \in \Pi, \gamma \neq \alpha, \gamma \neq \beta$ with three distinct neighbors. Then, using what we have just proved, Lemma 25, and Lemma 27, and interchanging α with β if necessary, we may find a sequence $\alpha_1 = \alpha, \alpha_2 = \beta, \alpha_3, \ldots, \alpha_n, \alpha_{n+1}$ of distinct elements of Π, such that α_1 and α_2 are double neighbors, α_i is a simple neighbor of α_{i-1} and α_{i+1} if $2 < i < n$, α_{n+1} is a simple neighbor of α_{n-1}, and no α_i has any neighbor α_j other than the ones just described. Put $u_1 = \mu_{\alpha_1}/|\mu_{\alpha_1}|$, $u_2 = (\mu_{\alpha_2} + \cdots + \mu_{\alpha_{n-1}})/|\mu_{\alpha_2} + \cdots + \mu_{\alpha_{n-1}}|$, $u_3 = \mu_{\alpha_n}/|\mu_{\alpha_n}|$, $u_4 = \mu_{\alpha_{n+1}}/|\mu_{\alpha_{n+1}}|$. This gives four unit vectors, linearly independent by Definition 17. By Lemma 23, u_1, u_3, and u_4 are orthogonal. We have $|\mu_{\alpha_2}|^2 = \cdots = |\mu_{\alpha_{n-1}}|^2$ and $(\mu_{\alpha_j}, \mu_{\alpha_{j+1}}) = -\frac{1}{2}|\mu_{\alpha_2}|^2$ for $2 < j < n-1$ by Lemma 23. Thus

$$|\mu_{\alpha_2} + \cdots + \mu_{\alpha_{n-1}}|^2 = |\mu_{\alpha_2}|^2 = |\mu_{\alpha_{n-1}}|^2,$$

and therefore $(u_2, u_1) = -\frac{1}{2}\sqrt{2}, (u_2, u_3) = -\frac{1}{2}, (u_2, u_4) = -\frac{1}{2}$. But then $1 = |u_2|^2 = (u_2, u_1)^2 + (u_2, u_3)^2 + (u_2, u_4)^2$, contradicting the linear independence of u_1, u_2, u_3, u_4. Q.E.D.

30. Corollary. Let Λ be simple. If any element of Π has a double neighbor, then the elements of Π may be enumerated in a sequence $\alpha_1, \alpha_2, \ldots, \alpha_m$, such that α_i has α_{i-1} and α_{i+1} as neighbors, but has no other neighbors.

Proof. This follows by Lemma 25, Lemma 27, and the preceding Lemma. Q.E.D.

31. Corollary. Let Λ be simple, let some element of Π have a double neighbor, and let Π be enumerated as in the preceding corollary. Let $m > 4$. Then either α_1 or α_m has a double neighbor.

Proof. If this is false then by the preceding corollary Π contains a subset Π_0 of 5 roots $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$, each β_i having β_{i-1} and β_{i+1} as its only neighbors in Π_0, such that β_3 is a double neighbor of β_4, and such that no β_i other than β_3 or β_4 has any neighbor which is not simple. Put

$$u = \frac{\mu_{\beta_1}}{|\mu_{\beta_1}|} + 2\,\frac{\mu_{\beta_2}}{|\mu_{\beta_2}|} + 3\,\frac{\mu_{\beta_3}}{|\mu_{\beta_3}|}\,; \quad v = \frac{\mu_{\beta_5}}{|\mu_{\beta_5}|} + 2\,\frac{\mu_{\beta_4}}{|\mu_{\beta_4}|}\,;$$

so that by Definition 17 the vectors u and v are linearly independent. Using Lemma 23 we find that

$$|u|^2 = 1^2 + 2^2 + 3^2 - 2 - 6 = 6$$

and similarly

$$|v|^2 = 3.$$

Using Lemma 23 once more, we find that

$$(u, v)^2 = 4 \cdot 9 \cdot \tfrac{1}{4}\, n_{\beta_3, \beta_4} = 18.$$

But then, by Schwarz' inequality, $18 = (u, v)^2 < |u|^2\,|v|^2 = 18$. This contradiction proves the present lemma. Q.E.D.

32. Corollary. Let Λ be simple, and let some element of Π have a double neighbor. Then the elements of Π may be enumerated in a sequence $\alpha_1, \ldots, \alpha_m$, such that α_i has α_{i-1} and α_{i+1} as neighbors, but has no other neighbors, and such that either

(a) α_1 has α_2 as a double neighbor, the doubling being on α_1 while no α_j other α_1 or α_2 has a double neighbor; or

(b) α_1 has α_2 as a double neighbor, the doubling being on α_2, while no α_j other than α_1 or α_2 has a double neighbor; or

(c) $n = 4$, α_2 has α_3 as a double neighbor, the doubling being on α_2, while no α_j other than α_2 and α_3 has a double neighbor.

This completes our analysis of the structure of a simple set of roots of a simple complex Lie algebra in case any root has a double (or triple) neighbor. We now turn to analyse the remaining case, in which every pair of neighboring roots is a pair of simple neighbors.

The following lemma is in close analogy with Lemma 29.

33. Lemma. Let Λ be simple. Let $\alpha \in \Pi$ have three distinct neighbors. Then no other element of Π has three distinct neighbors.

Proof. By Lemma 29, no element of Π has a double neighbor, and, by Lemma 26, no element has more than three neighbors. Now suppose that the present lemma is false. Then, using Lemmas 27 and 25, we may find a sequence $\beta, \alpha_1, \alpha_2 = \alpha, \ldots, \alpha_n, \gamma$ of distinct elements of Π, such that α_j has α_{j-1} and α_{j+1} as simple neighbors, $2 \leqslant j \leqslant n-1$, β has α_2 as a simple neighbor, γ has α_{n-1} as a simple neighbor, and such that no element of the set $\Pi_0 = \{\beta, \alpha_1, \ldots, \alpha_n, \gamma\}$ has any neighbor in Π_0 other than the ones just described. Put $u_1 = \mu_\beta/|\mu_\beta|$, $u_2 = \mu_{\alpha_1}/|\mu_{\alpha_1}|$,

$$u_3 = (\mu_{\alpha_2} + \cdots + \mu_{\alpha_{n+1}})/|\mu_{\alpha_2} + \cdots + \mu_{\alpha_{n-1}}|, u_4 = \mu_{\alpha_n}/|\mu_{\alpha_n}|, u_5 = \mu_\gamma/|\mu_\gamma|.$$

This gives five unit vectors, linearly independent by Definition 17. By Lemma 23, u_1, u_2, u_4, and u_5 are orthogonal. We have $|\mu_{\alpha_2}|^2 = \cdots = |\mu_{\alpha_{n-1}}|^2$ and $(\mu_{\alpha_j}, \mu_{\alpha_{j+1}}) = -\frac{1}{2}|\mu_{\alpha_2}|^2$ for $2 < j < n-1$ by Corollary 23. Thus

$$|\mu_{\alpha_2} + \cdots + \mu_{\alpha_{n-1}}|^2 = |\mu_{\alpha_2}|^2 = |\mu_{\alpha_{n-1}}|^2,$$

and therefore $(u_1, u_3) = (u_2, u_3) = (u_4, u_3) = (u_5, u_3) = -\frac{1}{2}$ by Lemma 23. But then $1 = |u_3|^2 = (u_1, u_3)^2 + (u_2, u_3)^2 + (u_4, u_3)^2 + (u_5, u_3)^2$, contradicting the linear independence of $u_1, \ldots, u_5$. This contradiction proves our lemma. Q.E.D.

The following lemma is in analogy with Corollary 31.

34. Lemma. Let Λ be simple, and suppose that no two neighboring elements of Π are either double or triple neighbors. Then Π contains

(a) no set Π_0 of seven roots $\alpha_1, \alpha_2, \ldots, \alpha_7$ such that α_1 neighbors α_2 neighbors α_3, α_4 neighbors α_5 neighbors α_3, α_6 neighbors α_7 neighbors α_3; and such that no α_j in Π_0 has any neighbor in Π_0 but the ones just described.

(b) no set Π_1 of nine roots $\beta_1, \ldots, \beta_8, \gamma$ such that β_j neighbors β_{j-1} and β_{j+1} for $2 \leqslant j \leqslant 7$, γ neighbors β_6, and such that no element of Π_1 has any neighbors in Π_1 but the ones just described.

(c) No set Π_2 of eight roots $\alpha_1, \ldots, \alpha_7$ and β such that α_j neighbors α_{j-1} and α_{j+1} for $2 \leqslant j \leqslant 6$, β neighbor α_4, and such that no element in Π_2 has any neighbors in Π_2 but the ones just described.

Proof. (a) If a set Π_0 as above exists, put

$$u_1 = \frac{\mu_{\alpha_1}}{|\mu_{\alpha_1}|} + 2\frac{\mu_{\alpha_2}}{|\mu_{\alpha_2}|}, \quad u_2 = \frac{\mu_{\alpha_4}}{|\mu_{\alpha_4}|} + 2\frac{\mu_{\alpha_5}}{|\mu_{\alpha_5}|},$$

$$u_3 = \frac{\mu_{\alpha_6}}{|\mu_{\alpha_6}|} + 2\frac{\mu_{\alpha_7}}{|\mu_{\alpha_7}|}, \quad u_4 = \frac{\mu_{\alpha_3}}{|\mu_{\alpha_3}|}.$$

(b) If a set Π_1 as above exists, put

$$u_1' = \frac{\mu_{\beta_1}}{|\mu_{\beta_1}|} + 2\frac{\mu_{\beta_2}}{|\mu_{\beta_2}|} + \cdots + 5\frac{\mu_{\beta_5}}{|\mu_{\beta_5}|}; \quad u_2' = \frac{\mu_{\beta_8}}{|\mu_{\beta_8}|} + 2\frac{\mu_{\beta_7}}{|\mu_{\beta_7}|};$$

$$u_3' = \frac{\mu_\gamma}{|\mu_\gamma|}; \quad u_4' = \frac{\mu_{\beta_6}}{|\mu_{\beta_6}|}.$$

Then $u_1, \ldots, u_4$ (resp. $u_1', \ldots, u_4'$) are 4 linearly independent vectors. The vectors u_4, u_3', u_4' are unit vectors; by Lemma 23, we have

$$|u_1|^2 = 3, \ |u_2|^2 = 3, \ |u_3|^2 = 3, \ |u_2'|^2 = 3,$$

$$|u_1'|^2 = 1^2 + \cdots + 5^2 - 1.2 - 2.3 - 3.4 - 4.5 = 15.$$

By Lemma 23, the vectors u_1, u_2, u_3 (resp. u_1', u_2', u_3') are mutually orthogonal. Moreover, by the same lemma

$$(u_1, u_4)^2 = 1, \ (u_2, u_4)^2 = 1, \ (u_3, u_4)^2 = 1;$$

$$(u_1', u_4')^2 = \tfrac{25}{4}, \ (u_2', u_4')^2 = 1, \ (u_3', u_4')^2 = \tfrac{1}{4}.$$

But then either the inequality

$$1 = |u_4|^2 < \frac{(u_1, u_4)^2}{|u_1|^2} + \frac{(u_2, u_4)^2}{|u_2|^2} + \frac{(u_3, u_4)^2}{|u_3|^2} = \frac{1}{3} + \frac{1}{3} + \frac{1}{3}$$

or

$$1 = |u_4'|^2 < \frac{(u_1', u_4')^2}{|u_1|^2} + \frac{(u_2', u_4')^2}{|u_2'|^2} + \frac{(u_3', u_4')^2}{|u_3|^2} = \frac{5}{12} + \frac{1}{3} + \frac{1}{4}$$

is contradicted.

(c) If a set Π_2 as above exists, put

$$u_1 = \frac{\mu_{\alpha_1}}{|\mu_{\alpha_1}|} + 2\frac{\mu_{\alpha_2}}{|\mu_{\alpha_2}|} + 3\frac{\mu_{\alpha_3}}{|\mu_{\alpha_3}|};$$

$$u_2 = \frac{\mu_{\alpha_7}}{|\mu_{\alpha_7}|} + 2\frac{\mu_{\alpha_6}}{|\mu_{\alpha_6}|} + 3\frac{\mu_{\alpha_5}}{|\mu_{\alpha_5}|};$$

$$u_3 = \frac{\mu_{\alpha_4}}{|\mu_{\alpha_4}|}; \quad u_4 = \frac{\mu_\beta}{|\mu_\beta|}.$$

The vectors u_3 and u_4 are unit vectors; the vectors u_1, u_2, u_3, u_4 are line-independent. By Lemma 23, the vectors u_1, u_2, and u_4 are orthogonal. Thus

$$(*)\ 1 = |u_3|^2 < \left(\frac{u_1}{|u_1|}, u_3\right)^2 + \left(\frac{u_2}{|u_2|}, u_3\right)^2 + (u_4, u_3)^2.$$

But, using Lemma 23 and Definition 24, we have

$$(u_1, u_3) = (u_2, u_3) = -\tfrac{1}{2}\cdot 3 = -\tfrac{3}{2}$$

$$(u_4, u_3) = -\tfrac{1}{2}$$

$$|u_1|^2 = |u_2|^2 = 1 + 4 + 9 - \tfrac{1}{2}\cdot 2\cdot(2 + 6) = 6.$$

Thus the right-hand side of inequality (*) above is $2 \cdot 6^{-1} \cdot \frac{9}{4} + \frac{1}{4} = 1$, giving a contradiction, and completing our proof. Q.E.D.

Lemma 34 concludes our analysis of the configuration of the vectors $\{\mu_\alpha\}$. We now wish to state a theorem summarizing the results of Lemmas 23–34. In doing this, we find it convenient to define a certain class of geometric symbols.

35. Definition. Let Λ be a semi-simple complex algebra, and let Π be a simple set of roots of Λ. Then the *Dynkin diagram* or *chromosome* of Λ

(relative to the set Π) is the linear graph Γ defined by the following three conditions:

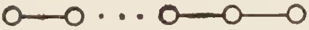
(a) the vertices of Γ are in 1 – 1 correspondence with the elements of Π.

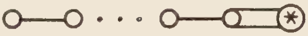
(b) two vertices of Γ are connected by a single line if and only if the corresponding elements of Π are simple neighbors; by a double line if and only if the corresponding elements of Π are double neighbors; by a triple line if and only if the corresponding elements of Π are triple neighbors.

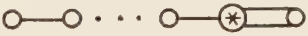
(c) a vertex of Γ is distinguished by an asterisk if and only if the corresponding element $\alpha \in \Pi$ has a double or triple neighbor β; the doubling or tripling being on α.

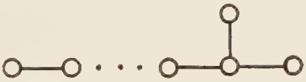
The following decisive theorem then summarizes the content of Lemmas 23–34.

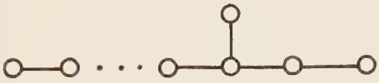
36. Theorem. Let A be a simple complex Lie algebra; let Π be a simple set of roots of A. Let Π have m points. Then the Dynkin diagram of A (relative to Π) is of one of the following types:

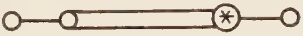
Type A_m:  $m \geq 1$; or

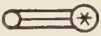
Type B_m:  $m \geq 2$; or

Type C_m:  $m \geq 3$; or

Type D_m:  $m \geq 4$; or

Type E_m:  $6 \leq m \leq 8$; or

Type F_4:  or

Type G_2: 

Two simple complex algebras with the same Dynkin diagram are isomorphic.

We shall see below that the converse of the final assertion of the above theorem is valid: If two simple complex Lie algebras are isomorphic, their Dynkin diagrams are the same. Thus, the type of the Dynkin diagram of A is a characteristic of A, and we may speak of a simple Lie algebra as an algebra of type A, B, E_8, F_4, G_2, etc. The four classes A, B, C, and D of simple Lie algebras are called the principal types of simple algebras; the five remaining algebras E_6, E_7, E_8, F_4 and G_2 are known as the *exceptional* simple algebras.

Not finally that while Theorem 36 gives a complete classification of simple complex Lie algebras, it does not assert that an algebra of each of the types which it describes actually exists. This additional assertion is true, however, and we turn now to the task of establishing it, and of describing the various types of simple algebras in more detail.

4. Models of the Simple Complex Algebras

Theorem 3.36 describes all possible simple complex Lie algebras. Our first aim in the present section will be to show that all the Lie algebras described by Theorem 3.36 actually exist. We will accomplish this aim by the pleasantly explicit method of displaying a specific model of each of the algebras. The following simple lemma will facilitate our computations.

1. Lemma. Let Λ be a complex Lie algebra, M a subalgebra of Λ such that $M^2 = 0$. Suppose that $\{\alpha\}$ is a set of roots of M in Λ, including the root zero, and that

(i) If $\alpha \in \{\alpha\}$, $-\alpha \in \{\alpha\}$.
(ii) The root spaces Λ_α, $\alpha \neq 0$, $\alpha \in \{\alpha\}$ are 1-dimensional.
(iii) If $0 \neq \lambda_\alpha \in \Lambda_\alpha$, $0 \neq \lambda_{-\alpha} \in \Lambda_{-\alpha}$, and $\alpha \neq 0$, it follows that $[\lambda_\alpha, \lambda_{-\alpha}] \neq 0$.
(iv) For each $\mu \in M$, there exists a $\mu' \in M$ such that $\alpha(\mu)$ and $\alpha(\mu')$ are complex conjugates for all $\alpha \in \{\alpha\}$.
(v) No non-zero $\mu \in M$ satisfies $\alpha(\mu) = 0$ for all $\alpha \in \{\alpha\}$.
(vi) $\Lambda = M \oplus \sum_{\substack{\alpha \in \{\alpha\} \\ \alpha \neq 0}} \Lambda_\alpha$.

Then Λ is semi-simple. Moreover, $\{\alpha\}$ is the set of all roots of M in Λ.

Proof. If the set A of roots of M in Λ contained any element other than those in the set $\{\alpha\}$, then $\Lambda = \sum_{\alpha \in A} \oplus \Lambda_\alpha$ would contradict (vi). This proves our final assertion.

For the rest, choose $0 \neq \lambda_\alpha \in \Lambda_\alpha$ for all $\alpha \in \{\alpha\}$, $\alpha \neq 0$.

Then, by hypothesis (vi), we may write any pair λ, λ' of Λ as $\lambda = \mu + \sum_{\substack{\alpha \in \{\alpha\} \\ \alpha \neq 0}} c_\alpha \lambda_\alpha$, $\lambda' = \mu' + \sum_{\substack{\alpha \in \{\alpha\} \\ \alpha \neq 0}} c'_\alpha \lambda_\alpha$, with complex coefficients c_α, c'_α, and with $\mu, \mu' \in M$. By Lemma 2.5, the operator $\text{Ad}\,\lambda_\alpha\, \text{Ad}\,\lambda_\beta$ maps Λ_γ into

$\Lambda_{\gamma+\alpha+\beta}$. Hence, the n-th power of this operator maps Λ_γ into $\Lambda_{\gamma+n(\alpha+\beta)}$. Since there are only finitely many roots, $(\mathrm{Ad}\,\lambda_\alpha\,\mathrm{Ad}\,\lambda_\beta)^n = 0$ for large n if $\alpha + \beta \neq 0$. Thus

$$(\lambda, \lambda') = \mathrm{tr}\,(\mathrm{Ad}\,\mu\,\mathrm{Ad}\,\mu') + \sum_{\substack{\alpha\in\{\alpha\}\\ \alpha\neq 0}} c_\alpha c'_\alpha\,\mathrm{tr}\,(\mathrm{Ad}\,\lambda_\alpha\,\mathrm{Ad}\,\lambda_{-\alpha}) \tag{1}$$

$$= \sum_{\alpha\in\{\alpha\}} (\alpha(\mu)\,\alpha(\mu') + \sum_{\substack{\alpha\in\{\alpha\}\\ \alpha\neq 0}} c_\alpha c'_\alpha\,\mathrm{tr}\,(\mathrm{Ad}\,\lambda_\alpha\,\mathrm{Ad}\,\lambda_{-\alpha}).$$

As we have already noted (cf. Corollary 2.16) $(\mathrm{Ad}\,(\nu)\,\lambda,\lambda') = -(\lambda, \mathrm{Ad}(\nu)\lambda')$ for $\nu, \lambda, \lambda' \in \Lambda$, and thus the set $\Lambda^{(0)}$ of λ such that $(\lambda, \lambda') = 0$ for all $\lambda' \in \Lambda$ is an ideal in Λ. By hypotheses (iv) and (v) and by (1) it follows that if $\Lambda^{(0)} \neq 0$, $\Lambda^{(0)}$ contains a pair λ_α, $\lambda_{-\alpha}$ of vectors. But then by hypothesis (iii) $[\lambda_\alpha, \lambda_{-\alpha}]$ is a non-zero element of $\Lambda^{(0)}$, which by Lemma 2.5 belongs to M; impossible by hypotheses (iv) and (v). This shows that the natural bilinear form of Λ is nonsingular: Hence, by Theorem 2.3, Λ is semi-simple, and our Lemma is proved. Q.E.D.

Now we turn to the examination of specific Lie algebras. The first concrete model we consider is the Lie algebra A_{n-1} of all $n \times n$ complex matrices of trace 0. We let e_{ij} denote the $n \times n$ matrix whose k, i-th element is $\delta_{ik}\delta_{jl}$. Then the matrices e_{ij}, $i \neq j$, $1 \leqslant i, j \leqslant n$, and $e_{ii} - e_{nn}$, $1 \leqslant i < n$, form a basis for A_{n-1}. We have $e_{ij}e_{kl} = \delta_{jk}e_{il}$. Let M denote the set of all diagonal matrices in A_{n-1}. Then $M^{(2)} = 0$. Define linear forms ω_i, $1 \leqslant i < n$, on M by putting $\omega_i(e_{ii} - e_{nn}) = 1, \omega_i(e_{jj} - e_{nn}) = 0$, $j \neq i$. Then, if $i \neq j$, the matrix e_{ij} is a root vector belonging to the root $\omega_i - \omega_j$ if $1 \leqslant i, j < n$; if $i \neq n$, the matrix e_{in} is a root vector belonging to the root $(\omega_i + \omega_1 + \cdots + \omega_{n-1})$; and if $i \neq n$ the matrix e_{ni} is a root vector belonging to the root $-(\omega_i + \omega_1 + \cdots + \omega_{n-1})$. Thus by Lemma 1, $\omega_j - \omega_i$, $i \neq j$, and $\pm(\omega_i + \omega_1 + \cdots + \omega_{n-1})$, $1 \leqslant i < n$ is the full set of roots and A_{n-1} is semi-simple. By Definition 3.17, $\omega_1 - \omega_2$, $\omega_2 - \omega_3, \ldots, \omega_{n-2} - \omega_{n-1}$, and $\omega_1 + \cdots + \omega_{n-2} + 2\omega_{n-1}$ form a simple set of roots. By Definition 3.24, each of these roots is a simple neighbor of the following root. Thus, by Lemma 3.27, A_{n-1} is a simple Lie algebra. Plainly, A_{n-1} is of type A_{n-1} in the sense of Theorem 3.36.

Next we consider the Lie algebra S_n of all $n \times n$ matrices skew-symmetric with respect to the non-singular symmetric bilinear form $x_1y_n + x_2y_{n-1} + \cdots + x_ny_1$ in n-dimensional vector space. The algebra S_n then consists of all matrices $\{C_{ij}\}$ satisfying $C_{n+1-i,n+1-j} = -C_{j,i}$, $1 \leqslant i, j \leqslant n$. Let M denote the set of all diagonal matrices in S_n, so that $M^{(2)} = 0$. Then, employing the notation e_{ij} as in the preceding paragraph, we see that the matrices $e_{ii} - e_{n+1-i,n+1-i}$, $1 \leqslant i \leqslant n/2$, form a basis for M. These matrices, together with the matrices $e_{ij} - e_{n+1-j,n+1-i}$, where $i \neq j$ and either

$$1 \leqslant i, j \leqslant \frac{n+1}{2}; \quad \text{or}$$

$$n \geqslant i > \frac{n+1}{2} \quad \text{and} \quad j < n+1-i; \quad \text{or}$$

$$n \geqslant j > \frac{n+1}{2} \quad \text{and} \quad i < n+1-j$$

form a basis for S_n. Define linear forms ω_i, $1 \leqslant i \leqslant n/2$, on M by putting $\omega_i(e_{ii} - e_{n+1-i,n+1-i}) = 1$, $\omega_i(e_{jj} - e_{n+1-j,n+1-j}) = 0$ for $i \neq j$ if $1 \leqslant i, j \leqslant n/2$. Then if $1 \leqslant i, j \leqslant n/2$, the matrix $e_{ij} - e_{n+1-j,n+1-i}$ belongs to the root $\omega_i - \omega_j$ of M in S_n. If $n \geqslant i > \dfrac{n+1}{2}$ and $j < n+1-i$, the matrix $e_{ij} - e_{n+1-j,n+1-i}$ belongs to the root $-(\omega_{n+1-i} + \omega_j)$ of M in S_n. If $n \geqslant j > \dfrac{n+1}{2}$ and $i < n+1-j$, the matrix $e_{i,j} - e_{n+1-j,n+1-i}$ belongs to the root $\omega_{n+1-j} + \omega_i$ of M in S_n. If n is even, those statements cover every matrix in the above-described basis for S_n. If n is odd, we have the additional matrices $e_{p,j} - e_{n+1-j,p}$, where $p = \dfrac{n+1}{2}$, and $1 \leqslant j < p$, and $e_{j,p} - e_{p,n+1-j}$ where again $p = \dfrac{n+1}{2}$ and $1 \leqslant j < p$, in the basis for S_n. The first set of matrices belong respectively to the roots $-\omega_j$ of M in S_n; the second to the roots ω_j of M in S_n. Thus, the non-zero roots of M in S_n are

n even; $n = 2m$: $\omega_i - \omega_j$ and $\pm(\omega_i + \omega_j)$, $1 \leqslant i, j \leqslant m, i \neq j$;

n odd; $n = 2m + 1$: $\omega_i - \omega_j$, $\pm(\omega_i + \omega_j)$, and $\pm\omega_i$, $1 \leqslant i, j \leqslant m$, $i \neq j$.

It follows by Lemma 1 that S_n is always semi-simple. Moreover, by Definition 3.17, we have the following simple set of roots:

$2m = n$ even: $\omega_1 - \omega_2, \omega_2 - \omega_3, \ldots, \omega_{m-1} - \omega_m, \omega_{m-1} + \omega_m$;

$2m + 1 = n$ odd: $\omega_1 - \omega_2, \omega_2 - \omega_3, \ldots, \omega_{m-1} - \omega_m, \omega_m$.

By Definition 3.24, each of the first $m - 1$ of these roots is a simple neighbor of the next. Moreover, by Definition 3.24

$n = 2m$ even: $\omega_{m-2} - \omega_{m-1}$ is a simple neighbor of $\omega_{m-1} + \omega_m$;

$n = 2m + 1$ odd: $\omega_{m-1} - \omega_m$ is a double neighbor of ω_m, the doubling being on ω_m.

Thus, by Lemma 3.27, S_{2m} and S_{2m+1} are simple. Plainly, S_{2m} is of type D_m and S_{2m+1} is of type B_m in the sense of Theorem 3.36.

Next we consider the Lie algebra Sp_{2m} of all $2m \times 2m$ matrices skew-symmetric with respect to the non-singular skew bilinear form $x_1 y_{2m} + \cdots + x_m y_{m+1} - x_{m+1} y_m - \cdots - x_{2m} y_1$. The algebra Sp_{2m} consists of all matrices $\{c_{ij}\}$ such that $c_{2m+1-i,2m+1-j} = -c_{ji}\varepsilon_i\varepsilon_j$, where $\varepsilon_i = 1$ if $1 \leqslant i \leqslant m$, $\varepsilon_i = -1$ if $m + 1 \leqslant i \leqslant 2m$. Let M denote the set of all diagonal matrices in S_n, so that $M^{(2)} = 0$. Then, employing the notation e_{ij} as in the preceding paragraphs, we see that the matrices $u_{ii} - e_{2m+1-i,2m+1-i}$, $1 \leqslant i \leqslant m$, form a basis for M. These matrices, together with the matrices

$e_{ij} - e_{2m+1-j,2m+1-i}$, where $i \neq j$ and $1 \leqslant i, j \leqslant m$

$e_{ij} + e_{2m+1-j,2m+1-i}$, where $2m \geqslant i \geqslant m + 1$ and $j \leqslant 2m + 1 - i$

$e_{ij} + e_{2m+1-j,2m+1-i}$, where $2m \geqslant j \geqslant m + 1$ and $i \leqslant 2m + 1 - j$

form a basis for Sp_{2m}. Define linear forms ω_i on M as in the above discussion of S_n $(n = 2m)$. Then, as above, the matrix $e_{ij} - e_{2m+1-j,2m+1-i}$, where $i \neq j$ and $1 \leqslant i, j \leqslant m$, belong to the roots $\omega_i - \omega_j$ of M in Sp_{2m}. The matrix $e_{ij} + e_{2m+1-j,2m+1-i}$, where $2m \geqslant i \geqslant m + 1$ and $j \leqslant 2m + 1 - i$, belongs to the root $-(\omega_{2m+1-i} + \omega_j)$ of M in SP_{2m}. The matrix $e_{ij} + e_{2m+1-j,2m+1-i}$, where $2m \geqslant j \geqslant m + 1$ and $i \leqslant 2m + 1 - j$, belongs to the root $\omega_{2m+1-j} + \omega_i$ of M in Sp_{2m}. Thus, the non-zero roots of M in Sp_{2m} are

$$\omega_i - \omega_j,\ i \neq j \quad \text{and} \quad \pm(\omega_i + \omega_j); \quad 1 \leqslant i, j \leqslant m.$$

It follows by Lemma 1 that Sp_{2m} is semi-simple. Moreover, by Definition 3.17, we have the following simple set of roots:

$$\omega_1 - \omega_2, \omega_2 - \omega_3, \ldots, \omega_{m-1} - \omega_m, 2\omega_m.$$

By Definition 3.24, each of the first $m - 1$ of these roots is a simple neighbor of the next. Moreover, $\omega_{m-1} - \omega_m$ is a double neighbor of $2\omega_m$, the doubling being on $\omega_{m-1} - \omega_m$. Thus, by Lemma 3.27, Sp_{2m} is simple. Plainly, Sp_{2m} is of type C_m in the sense of Theorem 3.36.

This completes our construction of a complex Lie algebra of each of the non-exceptional types. We turn now to the construction of a complex Lie algebra of each of the fixe exceptional types, using a method due to Freudenthal. We begin these constructions as follows. Let V be a finite dimensional complex vector space, and V' its conjugate space. Let Λ be a complex simple Lie algebra. Let ϱ be a representation of Λ on V; and let ϱ' be the representation of Λ on V' defined by

$$(\varrho'(\lambda)\, v')\,(v) = -v'(\varrho(\lambda)\, v), \quad \lambda \in \Lambda. \tag{1}$$

Then, for each $v' \in V'$ and $v \in V$, $v'(\varrho(\lambda)\, v)$ defines a linear functional on Λ. By Theorem 2.3 and Lemma 0.5, there exists a unique element $v' \cdot v \in \Lambda$ such that

$$v'(\varrho(\lambda)\, v) = (\lambda, v' \cdot v), \tag{2}$$

where (λ, λ') is the natural bilinear form in Λ. We suppose additionally that there exists a skew-symmetric trilinear form $\langle v_1, v_2, v_3\rangle$ in V, and a similar form in V', such that

$$\langle \exp(\varrho(\lambda))\, v_1, \exp(\varrho(\lambda))\, v_2, \exp(\varrho(\lambda))\, v_3\rangle = \langle v_1, v_2, v_3\rangle \tag{3}$$

$$\langle \exp(\varrho'(\lambda))\, v_1', \exp(\varrho'(\lambda))\, v_2', \exp(\varrho'(\lambda))\, v_3'\rangle = \langle v_1', v_2', v_3'\rangle$$

for $\lambda \in \Lambda$, $v_1, v_2, v_3 \in V$, $v_1', v_2', v_3' \in V'$. Thus, differentiating, we have

$$\langle \varrho(\lambda)\, v_1, v_2, v_3\rangle + \langle v_1, \varrho(\lambda)\, v_2, v_3\rangle + \langle v_1, v_2, \varrho(\lambda)\, v_3\rangle = 0 \tag{4}$$

$$\langle \varrho'(\lambda)\, v_1', v_2', v_3'\rangle + \langle v_1', \varrho'(\lambda)\, v_2', v_3'\rangle + \langle v_1', v_2', \varrho'(\lambda)\, v_3'\rangle = 0$$

for $\lambda \in \Lambda$, $v_1, v_2, v_3 \in V$, $v_1', v_2', v_3' \in V'$. The formulae

$$(v_1 \times v_2)\,(v_3) = \langle v_1, v_2, v_3\rangle$$

$$v_3(v_1' \times v_2') = \langle v_1', v_2', v_3'\rangle \tag{5}$$

define skew-symmetric bilinear mappings of $V \times V$ into V' and $V' \times V'$ into V respectively. We suppose finally that the following identity, which generalizes a familiar identity of vector analysis, holds:

$$(v_1 \times v_2)\,(v_1' \times v_2') = (v_1' \cdot v_2, v_2' \cdot v_1) - (v_2' \cdot v_2, v_1' \cdot v_1). \tag{6}$$

If all these hypotheses are satisfied, we may define a skew-symmetric bilinear product on the direct sum vector space $\Lambda \oplus V \oplus V'$, which turns this vector space into a Lie algebra, as follows. Since the product is to be bilinear we have only to define each of the products $[a, b]$, where $a \in \Lambda$ or V or V' and $b \in \Lambda$ or V or V', in such a way that the identity $[a, b] = -[b, a]$

and the Jacobi identity holds. We take the product $[\lambda, \lambda']$ of two elements of Λ over from the Lie-algebra structure of Λ, and put

$$[v, \lambda] = -[\lambda, v] = -\varrho(\lambda)\, v; \tag{7}$$

$$[v', \lambda] = -[\lambda, v'] = +\varrho'(\lambda)\, v';$$

$$[v_1, v_2] = v_1 \times v_2;\ [v_1', v_2'] = v_1' \times v_2'$$

$$[v, v'] = -[v', v] = -v' \cdot v$$

for $\lambda \in \Lambda$, $v_1, v_2, v \in V$, $v_1', v_2', v' \in V'$. The skew-symmetry of the product defined thereby is evident. In verifying the Jacobi identity $[[a, b], c] + [[b, c], a] + [[c, a], b] = 0$ we may suppose that $c \in \Lambda$ if any of $a, b, c \in \Lambda$. Then, if $b \in \Lambda$ also, the identity follows either because Λ is a Lie algebra or because ϱ is a representation. (And hence $-\varrho'(\lambda)$ is a representation!) If $b \in V$ or V' and $a \in V$ or V', the identity follows from an appropriate combination of (1), (2), (4), and (5). Thus, e.g., the expression (d, μ), where $\mu \in \Lambda$ and

$$d = [[v, v'], \lambda] + [[v', \lambda], v] + [[\lambda, v], v'], \text{ is equal to}$$

$$- ([v' \cdot v, \lambda], \mu) + (\varrho'(\lambda)\, v' \cdot v, \mu) - (v' \cdot \varrho(\lambda)\, v, \mu)$$

$$= +v'(\varrho([\mu, \lambda])\, v) + (\varrho'(\lambda)\, v')\,(\varrho(\mu)\, v) - v'(\varrho(\mu)\,\varrho(\lambda)\, v)$$

$$= +v'(\varrho([\mu, \lambda])\, v) + v'((\varrho(\lambda)\,\varrho(\mu) - \varrho(\mu)\,\varrho(\lambda))\, v) = 0$$

by (2) and the identity $(\mathrm{Ad}\,(\nu)\,\lambda, \mu) = -(\lambda, \mathrm{Ad}\,(\nu)\,\mu)$ valid for $\lambda, \mu, \nu \in \Lambda$. The only remaining cases of the Jacobi identity are then the cases $a, b \in V$, $c \in V'$, $a, b \in V'$, $c \in V$, $a, b, c \in V$, and $a, b, c \in V'$. In the two latter cases, the Jacobi identity follows by (5), (4), and (7). Thus, e.g., we have

$$[[v_1, v_2], v_3] + [[v_2, v_3], v_1] + [[v_3, v_1], v_2]$$

$$= +(v_1 \times v_2) \cdot v_3 + (v_2 \times v_3) \cdot v_1 + (v_3 \times v_1) \cdot v_2 .$$

If we call this expression d, and form (d, μ) with $\mu \in \Lambda$, then we get

$$+(v_1 \times v_2)\,(\varrho(\mu)\, v_3) + (v_2 \times v_3)\,(\varrho(\mu)\, v_1) + (v_3 \times v_1)\,(\varrho(\mu)\, v_2)$$

$$= \langle v_1, v_2, \varrho(\mu)\, v_3 \rangle + \langle \varrho(\mu)\, v_1, v_2, v_3 \rangle + \langle v_1, \varrho(\mu)\, v_2, v_3 \rangle = 0$$

by (5) and (4). In the two former cases, the Jacobi identity follows by (5), (2), and (6). Thus, e.g., we have

$$\begin{aligned}&[[v_1, v_2], v'] + [[v_2, v'], v_1] + [[v', v_1], v_2]\\ &= (v_1 \times v_2) \times v' - \varrho(v' \cdot v_2)\, v_1 + \varrho(v' \cdot v_1)\, v_2 .\end{aligned}$$

Applying a linear functional $u' \in V'$ to either side of this equation, we obtain

$$\begin{aligned}&\langle v_1 \times v_2, v', u'\rangle - u'(\varrho(v' \cdot v_2)\, v_1) + u'(\varrho(v' \cdot v_1)\, v_2)\\ &= (v_1 \times v_2)\,(v' \times u') - ((v' \cdot v_2), (u' \cdot v_1)) + ((v' \cdot v_1), (u' \cdot v_2))\\ &= 0\end{aligned}$$

by (5), (2), and (6). This verifies that, with the skew-symmetric product which has been defined, $\Lambda \oplus V \oplus V'$ becomes a Lie algebra.

To obtain a first example of a Lie algebra constructed in this way, we let Λ be the algebra of 3×3 matrices $\lambda = (\lambda^i_j)$ of trace zero, and put $(\varrho(\lambda)\, v)_i = \lambda^j_i v_j$ for each vector $v = (v_i)$ in 3-dimensional space V; here and in what follows we adopt the familiar "summation convention", according to which we sum over any index which occurs both as a superscript and as a subscript. We define the skew-symmetric triple product $\langle u, v, w\rangle$ of 3 vectors in the usual way by

$$\langle u, v, w > = \varepsilon^{ijk} u_i v_j w_k ,$$

where $\varepsilon^{ijk} = \varepsilon_{ijk} = 0$ if any two of i, j, k are the same, and where ε^{ijk} is the sign of the permutation $(\begin{smallmatrix}1 & 2 & 3\\ i & j & k\end{smallmatrix})$ otherwise. We put similarly

$$\langle u', v', w'\rangle = \tfrac{2}{27}\, \varepsilon_{ijk} u'^i v'^j w'^k$$

for $u' = (u'^i)$, $v' = (v'^i)$, $w' = (w'^i)$ in the space V'. Since $\exp(\varrho(\lambda))$ is a matrix of determinant 1 for each $\lambda \in \Lambda$, the identity (3) is easily verified, and the identity (4) follows. It is easily checked that the natural bilinear form in Λ, i.e., the trace of the mapping $\mu \to \operatorname{Ad}(\lambda_1)(\operatorname{Ad}(\lambda_2)(\mu))$, is $6 \operatorname{tr}(\lambda_1 \lambda_2)$. Thus the product $v' \cdot v$ of $v' \in V'$ and $v \in V$ is the matrix $\frac{1}{6}\{(v')^j v_i - \frac{1}{3}\, \delta^j_i (v')\,(v)\}$. We have therefore

$$(u' \cdot u, v' \cdot v) = \tfrac{1}{6} \cdot (u'(v)\, v'(u) - \tfrac{1}{3} \cdot u'(u)\, v'(v)) . \tag{8}$$

Let $u, v \in V$, $u', v' \in V'$. Since $\varepsilon^{ijk} \varepsilon_{lmn}$ is the sign of the permutation $(\begin{smallmatrix}i & j & k\\ l & m & n\end{smallmatrix})$, we have $\varepsilon^{ijk} \varepsilon_{ilm} = 3$ if $j = l$ and $k = m$, -3 if $j = m$ and $k = l$, and zero otherwise. Using this fact, we find using (8) that

$$\begin{aligned}(u \times v)\,(u' \times v') &= \tfrac{2}{9}\,(u'(u)\,v'(v) - u'(v)\,v'(u)) \\ &= (u' \cdot v, v' \cdot u) - (v' \cdot v, u' \cdot u).\end{aligned}$$

This verifies the identity (6); so that $\Lambda \oplus V \oplus V'$ is a Lie algebra. We analyzed the Lie algebra Λ earlier in the present section, and saw that the subalgebra M of diagonal matrices is a Cartan subalgebra of Λ. We may use this fact to analyze the algebra $\Lambda \oplus V \oplus V'$. In the usual way, we may regard each of Λ, V, and V' as being imbedded in $\Lambda \oplus V \oplus V'$. In particular, Λ is a subalgebra of $\Lambda \oplus V \oplus V'$.

As we saw in the paragraph immediately following the proof of Lemma 1, M has the basis $e_{11} - e_{33}, e_{22} - e_{33}$, and the roots $\pm(\omega_1 - \omega_2)$, $\pm(2\omega_1 + \omega_2)$, and $\pm(\omega_1 + 2\omega_2)$. If v_1, v_2, v_3 are the standard unit vectors $[1, 0, 0]$, $[0, 1, 0]$, and $[0, 0, 1]$ in V, then it is easily seen that v_i is a root vector of $M \subseteq \Lambda \oplus V \oplus V'$ belonging to the root ω_i if $i = 1$ or 2, and to the root $-(\omega_1 + \omega_2)$ if $i = 3$. Similarly, if v_1', v_2', v_3' are the standard unit vectors in V', then it is easily seen that v_i' is a root vector of M belonging to the root $-\omega_i$ if $i = 1$ or 2, and to the root $\omega_1 + \omega_2$ if $i = 3$.

Therefore the roots of M in $\Lambda \oplus V \oplus V'$ are $\pm\omega_i$, $\pm(\omega_1 - \omega_2)$, $\pm(\omega_1 + \omega_2)$, $\pm(2\omega_1 + \omega_2)$, and $\pm(\omega_1 + 2\omega_2)$. By Lemma 1, $\Lambda \oplus V \oplus V'$ is semi-simple. By Definition 3.17, $\omega_1 - \omega_2$ and ω_2 form a simple set of roots which by Definition 3.24 are triple neighbors, the tripling being on ω_2. Using Lemma 3.27 we see $\Lambda \oplus V \oplus V'$ is a simple algebra. This algebra is plainly of type G_2 in the sense of Theorem 3.36. It is 14-dimensional.

We may construct the 248-dimensional algebra of type E_8 in a very similar way. We now let Λ be the set of all 9×9 matrices $\lambda = (\lambda_j^i)$ of trace zero, so that Λ is 80-dimensional. We let V be the set of all tensors v_{ijk} of 3-indices, skew-symmetric in the three indices, so that V' is the set of all tensors v_{ijk}' of 3-indices, skew-symmetric in the three indices. Each of these spaces is 84-dimensional, so that the dimensions of $\Lambda \oplus V \oplus V'$ is 248. We define the skew-symmetric triple product of three elements of V by

$$\langle u\, v\, w \rangle = \varepsilon^{ijklmnopq} u_{ijk} v_{lmn} w_{opq},$$

where $\varepsilon^{ijklmnopq} = \varepsilon_{ijklmnopq}$ is zero if any two of its indices are the same, and is the sign of the permutation $\left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ i & j & k & l & m & n & o & p & q \end{smallmatrix}\right)$ otherwise. The skew-symmetric triple product of three elements of V' is defined similarly by

$$\langle u', v', w' \rangle = 2^{-5} 3^{-7} 7^{-1} \varepsilon_{ijklmnop} u'^{ijk} v'^{lmn} w'^{opq}.$$

We define the representation $\varrho(\lambda)$ of Λ on V in the manner customary in tensor analysis:

$$(\varrho(\lambda)\, v)_{ijk} = \lambda_i^l v_{ljk} + \lambda_j^l v_{ilk} + \lambda_k^l v_{ijl}.$$

It follows readily from the differential equation $\dfrac{d}{dt} \exp\,(t\varrho\,(\lambda)) = \varrho\,(\lambda) \exp\,(t\varrho\,(\lambda))$ that

$$(\exp\,(\varrho\,(\lambda))\, v)_{ijk} = a_i^l a_j^m a_k^n v_{lmn},$$

where a is a matrix of determinant one. From this, the identity (3) follows without difficulty, so that the identity (4) also follows. It is easily verified that the natural bilinear form in Λ, i.e., the trace of the mapping $\mu \to \mathrm{Ad}\,(\lambda_1)\,(\mathrm{Ad}\,(\lambda_2)\,\mu)$, is $18\ \mathrm{tr}\,(\lambda_1\lambda_2)$. Thus the product $v' \cdot v$ of $v' \in V'$ and $v \in V$ is the matrix

$$(v' \cdot v)_i^j = \tfrac{1}{18}\,(v'^{jlm} v_{ilm} - \tfrac{1}{9}\,\delta_i^j v'(v)).$$

We have therefore

$$(u' \cdot u, v' \cdot v) = \tfrac{1}{18}\,(u'^{jlm} u_{ilm} v'^{ino} v_{jno} - \tfrac{1}{9}\, u'(u)\, v'(v)) \tag{9}$$

if $u, v \in V$ and $u', v' \in V'$.

Since $\varepsilon^{ijklmnopq}\,\varepsilon_{i'j'k'l'm'n'o'p'q'}$ is the sign of the permutation $\left(\begin{smallmatrix} i & j & k & l & m & n & o & p & q \\ i' & j' & k' & l' & m' & n' & o' & p' & q' \end{smallmatrix}\right)$, we find that

$$(u \times v)\,(u' \times v') = 9 \cdot 8 \cdot 7\ \varepsilon^{lmnopq}_{l'm'n'o'p'q'} u_{lmn} v_{opq} u'^{l'm'n'} v'^{o'p'q'}\ ),$$

where $\varepsilon^{lmnoqq}_{l'm'n'o'p'q'}$ is

(a) zero unless both its unbarred and its barred indices constitute the same set of six distinct elements, and

(b) the sign of the permutation $\left(\begin{smallmatrix} l & m & n & o & p & q \\ l' & m' & n' & o' & p' & q' \end{smallmatrix}\right)$ if both its unbarred and its barred indices constitute the same set of six distinct elements.

Therefore

$$(u \times v)\,(u' \times v') = 9 \cdot 8 \cdot 7 \cdot 6(9 u_{ijk} v_{lmn} u'^{ijl} v'^{kmn} - 9 u_{ijk} v_{lmn}\, u'^{ilm} v'^{jkn}$$
$$+\, v'(u)\, u'(v) - u'(u)\, v'(v))$$

Using formula (9) above we find that

$$(u \times v)\,(u' \times v') = (9^2 \cdot 8 \cdot 7 \cdot 6 \cdot 18)\,(2^{-5} 3^{-7} 7^{-1})\,((u' \cdot u, v' \cdot v)$$
$$-(v' \cdot u, u' \cdot v)),$$

verifying formula (6). This shows that $\Lambda \oplus V \oplus V'$, with the skew-symmetric bilinear product which we have defined, is indeed a Lie algebra.

We analyzed the Lie algebra Λ early in the present section, and saw that the subalgebra M of diagonal matrices is a Cartan subalgebra of Λ. Using the notations of our earlier analysis, we recall that M has the basis $e_{11} - e_{99}, e_{22} - e_{99}, \ldots, e_{88} - e_{99}$, and the roots $\pm(\omega_i - \omega_j)$, where $\omega_i \neq \omega_j$, and $\pm(\omega_1 + \cdots + \omega_8 + \omega_i)$, $1 \leqslant i, j \leqslant 8$. In the usual way, we may regard each of Λ, V, and V' as being imbedded in $\Lambda \oplus V \oplus V'$. In particular, Λ is a subalgebra of $\Lambda \oplus V \oplus V'$. For each l, m, n such that $1 \leqslant l < m < n \leqslant 9$, let $v^{(l,m,n)}$ denote the skew-symmetric tensor whose element $v^{(lmn)}_{ijk}$ is zero unless l, m, n and i, j, k constitute the same set of 3 elements, and which is equal to the sign of the permutation $\left(\begin{smallmatrix} l & m & n \\ i & j & k \end{smallmatrix}\right)$ if l, m, n and i, j, k constitute the same set of 3 elements. Then $v^{(l,m,n)}$ is a root vector of $M \subseteq \Lambda + V + V'$ belonging to the root $\omega_l + \omega_m + \omega_n$; here we write $\omega_9 = -\omega_1 - \omega_2 \cdots -\omega_8$ for notational simplicity. Using the notation ω_9 in this way, and making a similar calculation for V', we see that the roots of M in $\Lambda \oplus V \oplus V'$ are $\omega_i - \omega_j$, where $1 \leqslant i, j \leqslant 9$, $i \neq j$ and $\pm(\omega_i + \omega_j + \omega_k)$, where $1 \leqslant i < j < k \leqslant 9$. By Lemma 1, $\Lambda \oplus V \oplus V'$ is semi-simple. By Definition 3.17, the roots $\omega_1 - \omega_2$, $\omega_2 - \omega_3, \ldots, \omega_7 - \omega_8$, and $\omega_6 + \omega_7 + \omega_8$ form a simple set of roots; by Definition 3.24, each of the first 7 roots is a simple neighbor of the next, and $\omega_5 - \omega_6$ is a simple neighbor of $\omega_6 + \omega_7 + \omega_8$. Using Lemma 3.27, we see that $\Lambda \oplus V \oplus V'$ is a simple algebra. This algebra is plainly of type E_8 in the sense of Theorem 3.36.

Having constructed a model of the Lie algebra of type E_8, we may use the following lemma to construct Lie algebras of types E_7 and E_6.

2. Lemma. Let Λ be a simple algebra, M a Cartan subalgebra, and let Π be a simple set of roots of Λ. For each $\alpha \in \Pi$, let $0 \neq \lambda_{\pm\alpha} \in \Lambda_{\pm\alpha}$. Let Π_0 be a subset of Π, and let $\Lambda^{(0)}$ be the subalgebra of Λ generated by the elements $\lambda_{\pm\alpha}$, $\alpha \in \Pi_0$. Then $\Lambda^{(0)}$ is semi-simple. Moreover, the Dynkin diagram Γ_0 of $\Lambda^{(0)}$ is obtained from the Dynkin diagram Γ of Λ by the following rules:

(a) erase all the vertices of Γ which correspond to elements of Π not lying in Π_0, together with all lines connecting the erased vertices to any other vertex;

(b) remove the asterisk from any vertex of Γ_0 whose (unique) double or triple neighbor has been erased.

Proof. Let M be the Cartan subalgebra of Λ, and let $\{\alpha\}$ be the set of roots of M in Λ. By Definition 3.17, we may write each $\beta \in \{\alpha\}$ as β

$= \sum_{\alpha \in \Pi} \eta_\alpha(\beta)\, \alpha$, where the integers η_α are either all positive or all negative. Let A_0 be the set of non-zero roots β for which the integers $\eta_\alpha(\beta)$ are zero whenever $\alpha \in \Pi_0$. By Lemma 3.11, Corollary 3.2, and Corollary 3.19 (ii), every space Λ_β, $\beta \in A_0$, and every vector μ_β, $\beta \in A_0$, belongs to $\Lambda^{(0)}$. On the other hand, it is plain from Lemma 3.11 that the subspace of Λ spanned by Λ_β, $\beta \in \Lambda_0$, and μ_β, $\beta \in A_0$, is a subalgebra of Λ. This shows that

$$\Lambda^{(0)} = \sum_{\beta \in A_0} \oplus \Lambda_\beta \oplus M^{(0)},$$

where $M^{(0)}$ is the subspace of M spanned by the vectors μ_β, $\beta \in A_0$. It is plain that if $\beta \in A_0$ and $\lambda_\beta \in \Lambda_\beta$, λ_β is a root vector of $M^{(0)}$ in $\Lambda^{(0)}$, belonging to the root $\beta | M^{(0)}$. By Corollary 3.8, the vectors η_γ, $\gamma \in \Pi_0$, span $M^{(0)}$. Since these vectors are linearly independent by Definition 3.17, they form a basis for $M(0)$. It is plain from Lemma 3.13 that each of the roots $B|M^{(0)}$ assumes real values on each of the basis vectors μ_γ; thus, for each $\mu \in M^{(0)}$, there exists a μ' such that $(\beta|M^{(0)})(\mu')$ is the complex conjugate of $(\beta|M^{(0)})(\mu)$ for each root $\beta|M^{(0)}$.

If $\beta_1, \beta_2 \in A_0$, and $\beta_1|M^{(0)} = \beta_2|M^{(0)}$, then we have $(\beta_1 - \beta_2)(\mu_\beta) = 0$ for each $\beta \in A_0$, so that by Corollary 3.8 $(\mu_{\beta_1} - \mu_{\beta_2}, \mu_\beta) = 0$ for each $\beta \in A_0$. By Lemma 3.16, it follows that $\beta_1 = \beta_2$. Thus each of the root spaces of $M^{(0)}$ in $\Lambda^{(0)}$ is one-dimensional. It follows by Lemma 1 that $\Lambda^{(0)}$ is semi-simple. Since the vectors μ_γ, $\gamma \in \Pi_0$, are a basis for $M^{(0)}$, it follows in the same way that the set $\{\gamma|M^{(0)}|\gamma \in \Pi_0\}$ of linear functionals is linearly independent. Thus, by Definition 3.17, this set is a simple set of roots of $\Lambda^{(0)}$. It is plain that for $\alpha, \beta \in \pi_0$ we have $\alpha + n\beta \in A_0$ if and only if $(\alpha|M^{(0)}) + n(\beta|M^{(0)})$ is a root of $M^{(0)}$ in $\Lambda^{(0)}$. Thus all the remaining conclusions of the present lemma follow by comparison with Definitions 3.24 and 3.25. Q.E.D.

3. Corollary. Let the hypotheses of the preceding lemma hold. Then

(a) The vectors μ_α, $\alpha \in \Pi_0$, are linearly independent, and the space $M^{(0)}$ which they span is a Cartan subalgebra of $\Lambda^{(0)}$.

(b) Let A_0 be the set of all roots of Λ which are linear combinations of the roots $\alpha \in \Pi_0$. Then the set of roots of $M^{(0)}$ in $\Lambda^{(0)}$ is $\{\alpha|M^{(0)}|\alpha \in A_0\}$. All these roots are distinct. The subset $\{\alpha|M^{(0)}|\alpha \in \Pi_0\}$ is a simple set of roots of $M^{(0)}$ in $\Lambda^{(0)}$, giving rise to the Dynkin diagram Γ_0 described in the preceding lemma.

(c) With A_0 and $M^{(0)}$ as above, $\Lambda^{(0)}$ is the direct sum of $M^{(0)}$ and of the various root spaces Λ_α for which $\alpha \in A_0$.

To apply the preceding Lemma and Corollary, let $\Lambda^{(1)} = \Lambda \oplus V \oplus V'$ be the algebra of type E_8 constructed above, and M its Cartan subalgebra, as constructed above. Let $\omega_1, \ldots, \omega_8$ be the linear functionals on M defined in the course of our discussion of $\Lambda^{(1)}$. By Lemma 2, the Lie subalgebra $\Lambda^{(2)}$ of $\Lambda^{(1)}$ generated by the elements $\lambda_{\pm\alpha}$, where $\lambda_{\pm\alpha} \in \Lambda^{(1)}_{\pm\alpha}$, and $\alpha \in \Pi^{(2)} = \{\omega_2 - \omega_3, \omega_3 - \omega_4, \ldots, \omega_7 - \omega_8, \omega_6 + \omega_7 + \omega_8\}$ is of type E_7 in the sense of Theorem 3.36. By Corollary 3, the space $M^{(2)}$ spanned by the vectors μ_α, $\alpha \in \Pi^{(2)}$, is a Cartan subalgebra of $\Lambda^{(2)}$, and the seven linear functionals $\alpha|M^{(2)}$, $\alpha \in \Pi^{(2)}$, form a simple set of roots of $M^{(2)}$ in $\Lambda^{(2)}$. It follows immediately that the seven linear functionals $\eta_i = \omega_{i+1}|M^{(2)}$, $1 \leqslant i \leqslant 7$, are linearly independent. Using Corollary 3 once more, we see that the roots of $M^{(2)}$ in $\Lambda^{(2)}$ are $\eta_i - \eta_j$, where $1 \leqslant i, j \leqslant 7$ and $i \neq j$; $\pm(\eta_i + \eta_j + \eta_k)$ where $1 \leqslant i < j < k \leqslant 7$, and $\pm(\eta + \eta_2 + \cdots + \eta_7 - \eta_i)$, where $1 \leqslant i \leqslant 7$. If we count these roots, and note that $M^{(2)} = \Lambda^{(2)}_0$ is 7-dimensional, we see that $\Lambda^{(2)}$ is of $7 + 42 + 70 + 14 = 133$ dimensions.

Using Corollary 3 and the description of the root spaces of the Cartan subalgebra M in the Lie algebra $\Lambda \oplus V \oplus V'$ of type E_8 given above, the reader will verify the following statements without difficulty: The Lie algebra $\Lambda^{(2)}$ is the direct sum $\Lambda^{(2)} = \tilde{\Lambda}^{(2)} \oplus V^{(2)} \oplus V'^{(2)}$, where $\tilde{\Lambda}^{(2)}$ is the set of all 9×9 matrices $\lambda = (\lambda_{ij})$ of trace zero such that $\lambda_{j9} = \lambda_{9j} = 0$ if $j \neq 9$, $\lambda_{j1} = \lambda_{1j} = 0$ if $j \neq 1$, and $\lambda_{11} = \lambda_{99}$; $V^{(2)}$ is the set of skew-symmetric tensors with components v_{ijk} which vanish whenever just one of i, j, k is 1 or 9; and $V'^{(2)}$ is the dual vector space of all skew-symmetric tensors with components $(v')^{ijk}$ which vanish whenever just one of i, j, k is 1 or 9. The Cartan subalgebra $M^{(2)}$ is the set of diagonal matrices in $\tilde{\Lambda}^{(2)}$.

Next let $\Lambda^{(3)}$ be the Lie subalgebra of $\Lambda^{(2)}$ generated by the elements $\lambda_{\pm\alpha}$, where $\lambda_{\pm\alpha} \in \Lambda^{(2)}_{\pm\alpha}$, and $\alpha \in \Pi^{(3)} = \{\eta_2 - \eta_3, \eta_3 - \eta_4, \ldots, \eta_6 - \eta_7, \eta_5 + \eta_6 + \eta_7\}$. By Lemma 2, the algebra $\Lambda^{(3)}$ is of type E_6 in the sense of Theorem 3.36. By Corollary 3, the space $M^{(3)}$ spanned by the vectors μ_α, $\alpha \in \Pi^{(3)}$, is a Cartan subalgebra of $\Lambda^{(3)}$, and the six linear functionals $\alpha|M^{(3)}$, $\alpha \in \Pi^{(3)}$, form a simple set of roots of $M^{(3)}$ in $\Lambda^{(3)}$. It follows immediately that the six linear functionals $\xi_i = \eta_{i+5}|M^{(3)}$, where $i = \pm 1, \pm 2, \pm 3$, are linearly independent (we choose our notations to emphasize a certain symmetry). Using Corollary 3 once more, we see that the roots of $M^{(3)}$ in $\Lambda^{(3)}$ are $\pm(\xi_i - \xi_j)$ where $i < j$ and $i, j = \pm 1, \pm 2, \pm 3$; $\pm(\xi_i + \xi_j + \xi_k)$ where $i < j < k$ and $i, j = \pm 1, \pm 2, \pm 3$; and $\pm(\xi_{-3} + \cdots + \xi_3 - \xi_i)$. Calculating as above, we find that the dimension of $\Lambda^{(3)}$ is 78.

Using Corollary 3 and the description of the root spaces in the Cartan subalgebra M in the Lie algebra $\Lambda \oplus V \oplus V'$ of type E_8 given above, the reader will verify the following statements without difficulty. The Lie algebra $\Lambda^{(3)}$ is the direct sum $\Lambda^{(3)} = \tilde{\Lambda}^{(3)} \oplus V^{(3)} \oplus V'^{(3)}$, where $\tilde{\Lambda}^{(3)}$ is the set of all 9×9 matrices $\lambda = (\lambda_j)$ of trace zero such that $\lambda_{j9} = \lambda_{9j} = 0$ if $j \neq 9$, $\lambda_{j1} = \lambda_{1j} = 0$ if $j \neq 1$, and $\lambda_{2j} = \lambda_{j2} = 0$ if $j \neq 2$, while $\lambda_{11} = \lambda_{22} = \lambda_{99}$; $V^{(3)}$ is the set of skew-symmetric tensors v with components v_{ijk} which vanish whenever one or two but not all three of i, j, k are 1, 2, or 9; and $V'^{(3)}$ is the dual vector space of all skew-symmetric tensors with components $(v')^{ijk}$ which vanish whenever one or two but not all three of i, j, k is 1, 2, or 9. The Cartan subalgebra $M^{(3)}$ is the set of all diagonal matrices in $\tilde{\Lambda}^{(3)}$.

The Dynkin diagram of type E_6 has a certain evident symmetry. We may use this fact to construct an algebra of the remaining exceptional type F_4, as follows. As noted above, $\Pi^{(3)} = \{\xi_{-3} - \xi_{-2}, \dots, \xi_2 - \xi_3, \xi_1 + \xi_2 + \xi_3\}$ is a simple set of roots of $\Lambda^{(3)}$. Define a linear transformation $\alpha \to \alpha'$ in the space M' of linear functionals on M by requiring

$$(\xi'_i - \xi'_{i+1}) = -(\xi_{-(i+1)} - \xi_{-i}), \quad i = \pm 1, \pm 2, \pm 3 \tag{10}$$

$$\xi'_1 + \xi'_2 + \xi'_3 = \xi_1 + \xi_2 + \xi_3 .$$

Then, by Corollary 3.22, it follows that if for each $\alpha \in \Pi \cup (-\Pi)$ we choose an element $\lambda_\alpha \in \Lambda^{(3)}_\alpha$ in such a way that $(\lambda_\alpha, \lambda_{-\alpha}) = 1$ if $\alpha \in \Pi$, then there exists an automorphism Φ of $\Lambda^{(3)}$ mapping λ_α into $\lambda_{\alpha'}$, $\alpha \in \Pi \cup (-\Pi)$. It is plain from Corollary 3.22 that $\Phi^2 = I$. Let $\Lambda^{(4)}$ be the Lie subalgebra of $\Lambda^{(3)}$ defined by $\Phi(\lambda) = \lambda$. We shall verify that $\Lambda^{(4)}$ is of type F_4. To verify this, we have only to carry out the following elementary calculations.

From (10), we see that the transformation $\lambda \to \alpha'$ acts on the roots of $M^{(3)}$ in $\Lambda^{(3)}$ as follows:

$$(\xi'_i - \xi'_j) = (\xi_{-j} - \xi_{-i}), \quad i \neq j, \quad i, j = \pm 1, \pm 2, \pm 3; \tag{11}$$

$\pm(\xi'_i + \xi'_j + \xi'_k) = \pm(\xi_p + \xi_q + \xi_r)$, where $\{p, q, r\}$ is the complement of $-i, -j, -k$ in the set $\{-3, -2, -1, 1, 2, 3\}$;

$$\pm(\xi'_{-3} + \cdots + \xi'_3) = \pm(\xi_{-3} + \cdots + \xi_3).$$

We divide the non-zero roots of $M^{(3)}$ into two classes: those for which $\alpha = \alpha'$, and those for which $\alpha \neq \alpha'$. Into the first class fall the six roots $\pm(\xi_{-i} - \xi_i)$, $i = 1, 2, 3$, the sixteen roots $\pm(\xi_{\pm 1} + \xi_{\pm 2} + \xi_{\pm 3})$, and the two roots $\pm(\xi_{-3} + \cdots + \xi_3)$, 24 roots in all; into the second class fall all the 48 other non-zero roots.

We have $\xi_{-2} - \xi_2 = (\xi_{-2} - \xi_{-1}) + (\xi_{-1} - \xi_1) + (\xi_1 - \xi_2)$, and thus by Lemma 3.11 $0 \neq [[\lambda_{\xi_{-2}-\xi_{-1}}, \lambda_{\xi_{-1}-\xi_1}], \lambda_{\xi_1-\xi_2}] \in \Lambda_{\xi_{-2}-\xi_2}$. It follows immediately that $\Phi|\Lambda_{\xi_{-2}-\xi_2} = I$. We may show in just the same way $\Phi|\Lambda_{\xi_i-\xi_j} = I$, $i = \pm 1, \pm 2, \pm 3$. Since $\xi_{-1} + \xi_2 + \xi_3 = (\xi_{-1} - \xi_1) + (\xi_1 + \xi_2 + \xi_3)$, we have similarly $[\lambda_{\xi_1+\xi_2+\xi_3}, \lambda_{\xi_{-1}-\xi_1}] \in \Lambda_{\xi_{-1}+\xi_2+\xi_3}$; and, by similar arguments, have $\Phi|\Lambda_{\pm(\xi_{\pm 1}+\xi_{\pm 2}+\xi_{\pm 3})} = I$. Then, since $[\lambda_{\xi_{-3}+\xi_{-2}+\xi_{-1}}, \lambda_{\xi_1+\xi_2+\xi_3}] \in \Lambda_{\xi_{-3}+\cdots+\xi_3}$, we see that $\Phi|\Lambda_\alpha = I$ for each non-zero root α such that $\alpha = \alpha'$. Using $\Phi^2 = I$, and using the fact that the automorphism Φ preserves the natural bilinear form of $\Lambda^{(3)}$, it follows that for each non-zero root α we may choose $\hat{\lambda}_\alpha \in \Lambda_\alpha$, in such a way that $\Phi(\hat{\lambda}_\alpha) = \hat{\lambda}_{\alpha'}$ and $(\hat{\lambda}_\alpha, \hat{\lambda}_{-\alpha}) = 1$.

A basis for $\Lambda^{(4)}$ then consists of

(i) a basis for

$$M^{(4)} = \{\mu \in M | \Phi(\mu) = \mu\}; \tag{12}$$

(ii) the 24 vectors $\hat{\lambda}_\alpha$, where $\alpha = \alpha'$ is a non-zero root of $M^{(3)}$ in $\Lambda^{(3)}$;
(iii) the 24 vectors $\hat{\lambda}_\alpha + \hat{\lambda}_{\alpha'}$ where $\alpha \neq \alpha'$ is a non-zero root of $M^{(3)}$ in $\Lambda^{(3)}$.

By Corollary 3.22 and Corollary 3.8, $\Phi(\mu_\alpha) = \mu_{\alpha'}$ for $\alpha \in \Pi$, so that, noting that two of the elements of Π satisfy $\alpha = \alpha'$ and that four satisfy $\alpha \neq \alpha'$, we see that $M^{(4)}$ is 4-dimensional. Thus $\Lambda^{(4)}$ is $4 + 24 + 24 = 52$-dimensional.

Since $\Phi(\mu_\alpha) = \mu_{\alpha'}$ for $\alpha \in \Pi$, it follows by Corollary 3.8 that $\alpha|M^{(4)} = \alpha'|M^{(4)}$ for $\alpha \in \Pi$, and thus for every linear functional α on $M^{(3)}$. It is plain that $[M^{(4)}, M^{(4)}] = 0$, that each of the vectors $\hat{\lambda}_\alpha$ described above under (ii) is a root vector of $M^{(4)}$ in $\Lambda^{(4)}$ belonging to the root $\alpha|M^{(4)}$, and that each of the vectors $\hat{\lambda}_\alpha + \hat{\lambda}_{\alpha'}$ described above is a root vector of $M^{(4)}$ in $\Lambda^{(4)}$ belonging to the root $\alpha|M^{(4)}$. All these roots take on real values on each of the four basic vectors $\mu_{\xi_{-3}-\xi_{-2}} + \mu_{\xi_2-\xi_3}$, $\mu_{\xi_{-2}-\xi_{-1}} + \mu_{\xi_1-\xi_2}$, $\mu_{\xi_1-\xi_{-1}}$, $\mu_{\xi_1+\xi_2+\xi_3}$ of $M^{(4)}$. Thus, for each $\lambda \in \Lambda^{(4)}$, there exists a $\tilde{\lambda} \in \Lambda^{(4)}$ such that $\beta(\tilde{\lambda})$ is the complex conjugate of $\beta(\lambda)$, β here denoting any arbitrary root $\beta = \alpha|M^{(4)}$ of $M^{(4)}$ in $\Lambda^{(4)}$.

If $\alpha|M^{(4)} = \beta|M^{(4)}$, then $(\mu_\alpha, \mu + \Phi(\mu)) = (\mu_\beta, \mu + \Phi(\mu))$ for each $\mu \in M^{(3)}$, and thus, by Corollary 3.8, $\alpha + \alpha' = \beta + \beta'$. Inspection of the equations (11) shows that this can only happen if $\alpha = \beta'$, or $\alpha = \beta$. Thus the root spaces of $M^{(4)}$ in $\Lambda^{(4)}$ are 1-dimensional. It follows by Lemma 1 that $\Lambda^{(4)}$ is semi-simple and that the non-zero roots of $M^{(4)}$ in $\Lambda^{(4)}$ are the linear forms $\frac{1}{2}(\alpha + \alpha')$, where α denotes an arbitrary non-zero root of $M^{(3)}$ in $\Lambda^{(3)}$. Thus, by Definition 3.17, and using the fact that $M^{(4)}$ is four-dimensional, the roots $\frac{1}{2}(\xi_{-3} - \xi_{-2} + \xi_2 - \xi_3)|M^{(4)}$, $\frac{1}{2}(\xi_{-2} - \xi_{-1} + \xi_1$

$- \xi_2)|M^{(4)}$, $(\xi_{-1} - \xi_1)|M^{(4)}$, and $(\xi_1 + \xi_2 + \xi_3)|M^{(4)}$ form a simple set of roots. The first of these roots is a neighbor of the second, the third is a double neighbor of the second, the doubling being on the third, and the third is a neighbor of the fourth. Thus, by Lemma 3.27 $\Lambda^{(4)}$ is simple. By Theorem 3.36, $\Lambda^{(4)}$ must be of type F_4, and we have constructed an algebra of every possible type.

A more explicit account of the roots of $\Lambda^{(4)}$ will be of use to us later, for which reason we derive it now. Let us introduce linear functionals $\omega_1, \omega_2, \omega_3, \omega_4$ on $M^{(4)}$ in such a way that the four linear functionals $\frac{1}{2}(\xi_{-3} - \xi_{-2} + \xi_2 - \xi_3)|M^{(4)}$, $\frac{1}{2}(\xi_{-2} - \xi_{-1} + \xi_1 - \xi_2)|M^{(4)}$, $(\xi_{-1} - \xi_1)|M^{(4)}$, and $(\xi_1 + \xi_2 + \xi_3)|M^{(4)}$ have the form $-\frac{1}{2}(\omega_1 + \omega_2 + \omega_3 + \omega_4)$, ω_3, $\omega_2 - \omega_3$, $\omega_1 - \omega_2$. Then it is easily verified that the restrictions to $M^{(4)}$ of the six roots $\pm(\xi_{-j} - \xi_j)$, $j = 1, 2, 3$ have the form $\pm(\omega_2 - \omega_3)$, $\pm(\omega_2 + \omega_3)$, $\pm(\omega_1 + \omega_4)$; the restrictions to $M^{(4)}$ of the sixteen roots $\pm(\xi_{\pm 1} + \xi_{\pm 2} + \xi_{\pm 3})$ have the form $\pm(\omega_1 - \omega_2)$, $\pm(\omega_1 \pm \omega_3)$, $\pm(\omega_4 + \omega_2)$, $\pm(\omega_1 + \omega_2)$, $\pm(\omega_3 - \omega_4)$, $\pm(\omega_3 + \omega_4)$, $\pm(\omega_4 - \omega_2)$; and the restrictions to $M^{(4)}$ of the two roots $\pm(\xi_{-3} + \cdots + \xi_3)$ have the form $\pm(\omega_1 - \omega_4)$. The remaining roots are the restrictions to $M^{(4)}$ of the twelve linear functionals $\pm\frac{1}{2}(\xi_i - \xi_{-i}) \pm \frac{1}{2}(\xi_j - \xi_{-j})$, $i \neq j$, $i, j = 1, 2, 3$, and of the twelve linear functionals $\pm\frac{1}{2}(\xi_i + \xi_{-i} + \xi_j + \xi_{-j} + 2\xi_{\pm k})$, where no two of i, j, and k are equal, and $1 \leqslant i, j, k \leqslant 3$. Expressing these roots in terms of the linear functionals $\omega_1, \omega_2, \omega_3, \omega_4$, we verify easily that the 48 non-zero roots of $M^{(4)}$ in $\Lambda^{(4)}$ have the form $\pm\omega_i \pm \omega_j$, where $1 \leqslant i, j \leqslant 4$ and $i \neq j$; $\pm\omega_i$, $1 \leqslant i \leqslant 4$; and $\frac{1}{2}(\pm\omega_1 \pm \omega_2 \pm \omega_3 \pm \omega_4)$.

The following theorem summarizes the work of the present section, adding a remark on the dimensions of the non-exceptional simple complex algebras which the reader will easily verify.

4. Theorem. There exists a simple complex Lie algebra of each of the types described in Theorem 3.36. Moreover,

(a) The algebra of type A_m is of dimension $m(m + 2)$, $m \geqslant 1$;
(b) The algebra of type B_m is of dimension $m(2m + 1)$, $m \geqslant 2$;
(c) The algebra of type C_m is of dimension $m(2m + 1)$, $m \geqslant 3$;
(d) The algebra of type D_m is of dimension $m(2m - 1)$, $m \geqslant 4$;
(e) The algebra of type E_8 is of dimension 248;
(f) The algebra of type E_7 is of dimension 133;
(g) The algebra of type E_6 is of dimension 78;
(h) The algebra of type F_4 is of dimension 52;
(i) The algebra of type G_2 is of dimension 14.

A system of roots for each of the simple complex Lie algebras is derived in the preceding paragraphs.

For later use, we note the following simple consequence of the constructions of the present section.

5. Theorem. Let Λ be a simple Lie algebra. Then we may choose a basis $\{\lambda_i\}$ for Λ in such a way that $[\lambda_i, \lambda_j] = \sum_k c^k_{ij} \lambda_k$, where all the coefficients c^k_{ji} are rational numbers. More precisely, such a basis may be constructed so as to consist of a basis for a certain Cartan subalgebra M of Λ and of a set of vectors λ_α which are elements of the various root spaces of Λ belonging to the non-zero roots of M in Λ.

Proof. For the types A_m, B_m, C_m, and D_m, this conclusion emerges immediately from the models we have constructed. Types G_2 and E_8 are covered by the same remark; we note explicitly that as a basis for the Cartan subalgebra M of E_8 to which the present theorem refers, we may choose the vectors $\mu_\alpha, \alpha \in \Pi$; Π being a simple set of roots of M in E_8. Types E_7 and E_6 were constructed from E_8 by the procedure described in Lemma 2 and Corollary 3; by Corollary 3, we get a basis for each of these algebras by choosing an appropriate subset of a basis of E_8 of the kind described by the present theorem. In particular, a basis for a Cartan subalgebra of E_7 and E_6 is obtained by choosing a subset of the vectors μ_α, $\alpha \in \Pi$. Thus our theorem is verified for types E_7 and E_6. Next we consider the type F_4. Let α' be as in the paragraph preceding formula (10). We constructed F_4 above as the set of all elements of E_6 satisfying $\Phi\lambda = \lambda$, where Φ was an automorphism of E_6 satisfying $\Phi\lambda_\alpha = \lambda_{-\alpha}, \alpha \in \Pi \cup (-\Pi)$, λ_α being vectors like those of the present theorem, as applied to E_6. It is then apparent from the present theorem (as applied to E_6) and from Lemma 3.11 and Corollary 3.19 (ii) that $\Phi\lambda_\alpha = r_\alpha\lambda_{\alpha'}$ for each non-zero root α, where r_α is a rational number. Thus the vectors $\hat{\lambda}_\alpha$ occuring in (12) may be taken to be rational multiples of the vectors λ_α. As a basis for the Cartan subalgebra $M^{(4)}$ (cf. (12)) we can take any four linearly independent vectors $[\lambda_\alpha, \lambda_{-\alpha}]$, for which $\alpha = \alpha'$. These are, of course, rational linear combinations of the six basic elements of the Cartan subalgebra of E_6 described by the present theorem. Thus the basis (12) of F_4 has all the properties described by the present theorem, and the present theorem is verified in all cases. Q.E.D.

5. Isomorphisms and Automorphisms

Our first step in analyzing the structure of a complex semi-simple Lie algebra Λ was to introduce a Cartan subalgebra M of Λ; every further set of objects connected with Λ, as e.g., the set of roots of M in Λ, was then defined in terms of M. Since Λ has many Cartan subalgebras, one might fear that the introduction of M introduces into our structural analysis an irrelevant arbitrariness. The following theorem shows that this is not so.

1. Theorem. Let Λ be a complex Lie algebra, and let M_1 and M_2 be Cartan subalgebras of Λ. Then there exists an automorphism σ of Λ such that $\sigma(M_1) = M_2$.

Our proof of Theorem 1 will show, in fact, that the automorphism σ can be taken to be inner in the sense of Definition 3 below.

To the proof of Theorem 1, we preface the following simple lemma and definition (cf. I.3.13 and I.5.3).

2. Lemma. Let Λ be a Lie algebra, and $\lambda \in \Lambda$. Then the mapping $\exp(\mathrm{Ad}(\lambda))$ is an automorphism of Λ.

Proof. If $\mu, \nu \in \Lambda$, the quantities $a_1(t) = \{\exp(t\,\mathrm{Ad}(\lambda))\}[\mu, \nu]$ and $a_2(t) = [\{\exp t\,\mathrm{Ad}(\lambda))\}\mu, \{\exp t\,\mathrm{Ad}(\lambda)\}\nu]$ are easily seen to satisfy the same linear differential equation $\frac{d}{dt} a_i(t) = \{\exp \mathrm{Ad}(\lambda)\} a_i(t)$, $i = 1, 2$; thus since $a_1(0) = a_2(0)$, our lemma follows. Q.E.D.

3. Definition. If Λ is a Lie algebra, the group of automorphisms of Λ generated by the automorphisms of the form $\exp(\mathrm{Ad}(\lambda))$, $\lambda \in \Lambda$, is called the *inner automorphism group* of Λ, and each of its elements is called an inner automorphism.

Now we give the proof of Theorem 1.

Proof of Theorem 1. Let $\tilde{\Lambda} = \{\lambda \in \Lambda |\ \mathrm{Ad}(\lambda)$ has the maximum possible number of eigenvalues$\}$. We saw in Lemma 2.8 that if $\lambda \in \tilde{\Lambda}$, then the set is a

$$M(\lambda) = \{\mu \in \Lambda | \mathrm{Ad}(\lambda)^n \mu = 0 \text{ for } n \text{ sufficiently large}\} \tag{1}$$

Cartan subalgebra of Λ. Suppose, conversely, that M is a Cartan subalgebra of Λ. Then, by Theorem 1.10, Λ decomposes as $\Lambda = M \oplus \sum \oplus \Lambda_\alpha$, where Λ_α are the root spaces belonging to the non-zero roots α of M in Λ. If we choose $\mu \in M$ such that $\alpha(\mu) \neq 0$ for each non-zero root, then it is plain that $M = M(\mu)$ in the sense of (1). Hence every Cartan subalgebra is of the form $M = M(\mu)$. Thus, if we let

$$\Lambda = \{\lambda \in \Lambda | M(\lambda) \text{ is a Cartan subalgebra}\},$$

then $\hat{\Lambda} \supseteq \dot{\Lambda}$, and every Cartan subalgebra is of the form $M(\lambda)$ with $\lambda \in \hat{\Lambda}$.

If λ_1 and λ_2 lie in $\hat{\Lambda}$, write $\lambda_1 \sim \lambda_2$ if there exists an inner automorphism σ such that $\sigma(M(\lambda_1)) = M(\lambda_2)$. We shall show that if $\lambda \in \hat{\Lambda}$, then every λ_1 sufficiently close to λ satisfies $\lambda \sim \lambda_1$. To this end, note that

$$\frac{d}{dt} \exp(\mathrm{Ad}\,(t\mu)) \times \lambda|_{t=0} = [\mu, \lambda];$$

on the other hand, $\frac{d}{dt}(\lambda + t\nu)|_{t=0} = \nu$. Since Λ is the sum of the range of Ad (λ) and the eigenspace of this operator belonging to the eigenvalue zero, we have $\Lambda = M(\lambda) + [\lambda, \Lambda]$. Thus, by the implicit function theorem (cf. Lemma 0.37), it follows that there exists a neighborhood V of λ in Λ such that each $\lambda_1 \in V$ can be written as $\lambda_1 = \sigma(\lambda_2)$, where σ is an inner automorphism, and $\lambda_2 \in M(\lambda)$. By Lemma 0.27, we may choose V sufficiently small so that $M(\lambda_2)$ and $M(\lambda)$ have the same dimension. But since $\lambda_2 \in M(\lambda)$ and $M(\lambda)$ is nilpotent, $M(\lambda) \subseteq M(\lambda_2)$. Thus $M(\lambda) = M(\lambda_2)$, so $\sigma(M(\lambda)) = \sigma(M(\lambda_2)) = M(\sigma\lambda_2) = M(\lambda_1)$, and therefore $\lambda_2 \sim \lambda$.

The set $\dot{\Lambda}$ is plainly open. If $\lambda \in \dot{\Lambda}$ and $\mu \in \Lambda$, then, by Lemma 0.28, the eigenvalues of Ad $(z\lambda + (1 - z)\,\mu)$ can be expressed as fractional power series in $(z - z_0)$ in the vicinity of any complex number z_0. Hence the number of eigenvalues changes only at isolated points. If we draw a curve from $z = 0$ to $z = 1$ not passing through any of these points, we see that $\dot{\Lambda}$ is dense. For the same reason, $\dot{\Lambda}$ is connected. Thus, since $\hat{\Lambda} \supseteq \dot{\Lambda}$, $\hat{\Lambda}$ is connected also. If we divide $\hat{\Lambda}$ into equivalence classes by the relation $\lambda_1 \sim \lambda_2$, then by what has been proved above, every equivalence class is open. Hence there can exist at most one equivalence class, and Theorem 1 is proved. Q.E.D.

Throughout the remainder of the present section, we let Λ *be a semisimple Lie algebra,* M *a Cartan subalgebra,* $\{\alpha\}$ *the set of roots of* M *in* Λ*, and* Π *a simple set of roots. For each* $0 \neq \alpha \in \{\alpha\}$*, we let* Λ_α *be the corresponding root space, and choose* $\lambda_\alpha \in \Lambda_\alpha$ *in such a way that* $(\lambda_\alpha, \lambda_{-\alpha}) = 1$ (cf. Corollary 3.2 and 3.7) *so that the vector* μ_α *of Corollary 3.8 is given by* $[\lambda_\alpha, \lambda_{-\alpha}] = \mu_\alpha$.

The structural analysis of Λ developed in section 3 involves the arbitrary choice of a simple set of roots. Here again it might be feared that an essential element of irrelevant arbitrariness is introduced. The following theorem shows that this is not so.

4. Theorem. Let Π' be a second simple set of roots of M in Λ. Then there exists an inner automorphism ϕ of Λ mapping M into itself, and such that the permutation $\alpha \to \alpha'$ of the roots of M in Λ defined by $\mu_{\alpha'} = \phi(\mu_\alpha)$ maps Π onto Π'.

5. Corollary. Any two simple sets of roots of M in Λ have the same Dynkin diagram.

6. Corollary. If two semi-simple Lie algebras are isomorphic, their Dynkin diagrams, formed with arbitrarily chosen Cartan subalgebras and arbitrarily chosen simple sets of roots of these Cartan subalgebras, are the same.

Proof of Corollaries. Let the two simple sets of roots be Π and Π'. Let ϕ be the automorphism whose existence is asserted in Theorem 4. By Corollary 3.8 and Definition 3.24, if $\alpha, \beta \in \Pi$ then α is a neighbor, double neighbor, or triple neighbor of β if and only if α' is a neighbor, double neighbor, or triple neighbor of β'; and a doubling or tripling lies on α if and only if a doubling or tripling lies on α'. This proves Corollary 5. Corollary 6 follows in an evident way from Corollary 5 and Theorem 1. Q.E.D.

To the proof of Theorem 4, we preface the following lemmas and definition.

7. Lemma. If α and β are non-zero elements of $|\alpha|$, so is

$$W_\alpha\beta = \beta - \frac{2\beta(\mu_\alpha)}{\alpha(\mu_\alpha)}\alpha,$$

Proof. By Corollary 3.12, we have $\dfrac{\beta(\mu_\alpha)}{\alpha(\mu_\alpha)} = -\frac{1}{2}(r_- + r_+)$, where r_+ and r_- are the greatest and the least r for which $\beta + r\alpha$ is a root. By the same corollary, $\beta + r\alpha$ is a root for all $r_- \leqslant r \leqslant r_+$. Since $r_+ \geqslant 0$, $2\dfrac{\beta(\mu_\alpha)}{\alpha(\mu_\alpha)} \leqslant -r_-$, and $W_\alpha\beta$ is a root. Since W_α maps α into $-\alpha$ and maps each linear functional β satisfying $\beta(\mu_\alpha) = 0$ onto itself, it is a reflection in a hyperplane in the space of linear functionals in M, and thus $W_\alpha\beta \neq 0$. Q.E.D.

8. Definition. The group of linear transformations of the space of linear functionals on M generated by the reflections W_α of Lemma 7 is called the *Weyl group* (of M in Λ). The maps W_α are called *Weyl reflections*.

Since by Lemma 7 the Weyl group permutes a finite set of linearly independent functionals (cf. Lemma 3.4) *among themselves, it is a finite group.*

9. Lemma. Let $\alpha \in \{\alpha\}, \alpha \neq 0$. Then there exists an inner automorphism ϕ of Λ, mapping M into itself, and such that $\phi(\mu_\beta) = \mu_{W_\alpha \beta}$ for each $\beta \in \{\alpha\}$, $\beta \neq 0$.

Proof. Write $\xi_0 = \lambda_\alpha + \lambda_{-\alpha}, \eta_0 = \lambda_{-\alpha} - \lambda_\alpha$. Then $\mathrm{Ad}\,(\xi_0)\,\mu_\alpha = \alpha(\mu_\alpha)\,\eta_0$, and $\mathrm{Ad}\,(\xi_0)\,\eta_0 = 2\mu_\alpha$. Consequently, if we put $\xi = c\xi_0, \eta = c\eta_0$, where $2c^2 = \alpha(\mu_\alpha)$, we have $\mathrm{Ad}\,(\xi)\,\mu_\alpha = \alpha(\mu_\alpha)\,\eta$, $\mathrm{Ad}\,(\xi)\,\eta = \alpha(\mu_\alpha)\,\mu_\alpha$. The exponential power series then gives $\exp\,(it\,\mathrm{Ad}\,(\xi))\,\xi = \xi$; $\exp\,(it\,\mathrm{Ad}\,(\xi))\,\eta = \cos\,(t\alpha\,(\mu_\alpha))\,\eta + i\,\sin\,(t\alpha\,(\mu_\alpha))\,\mu_\alpha$; $\exp\,(it\,\mathrm{Ad}\,(\xi))\,\mu_\alpha = \cos\,(t\alpha\,(\mu_\alpha))\,\mu_\alpha + i\,\sin\,(t\alpha(\mu_\alpha))\,\eta$. If we put $t = \pi(\alpha(\mu_\alpha))^{-1}$ and $\exp\,(it\,\mathrm{Ad}\,(\xi)) = \phi$, we have $\phi\mu_\alpha = -\mu_\alpha$. On the other hand, if $\mu \in M$ and $\alpha(\mu) = 0$, then $\mathrm{Ad}\,(\xi)\,\mu = 0$, and $\phi\mu = \mu$. Thus $\phi M = M$. Since the automorphism ϕ preserves the natural bilinear form of M, ϕ must be the reflection

$$\phi\mu = \mu - \frac{2(\mu_\alpha, \mu)}{(\mu_\alpha, \mu_\alpha)}\,\mu_\alpha, \; \mu \in M,$$

and thus our lemma is proved. Q.E.D.

Remark. We note, for future use, that the inner automorphism ϕ of the preceding lemma is

$$\exp\,(i\pi\,(2\alpha\,(\mu_\alpha))^{-1/2}\,\mathrm{Ad}\,(\lambda_\alpha + \lambda_{-\alpha})).$$

Now we can prove Theorem 4.

Proof of Theorem 4. Let $P(\Pi)$ be the set of non-zero roots β of M in Λ for which $\beta = \sum_{\alpha \in \Pi} n_\alpha \alpha$, with all $n_\alpha \geqslant 0$ (cf. Definition 3.17). Let $\alpha \in \Pi$. Note first that if $\beta \in P(\Pi)$ and $\beta \neq \alpha$, $W_\alpha \beta \in P(\Pi)$. Indeed, by Corollary 3.12, $W_\alpha \beta = \beta + r_+ \alpha$ where $r_+ \geqslant 0$ unless $\beta - \alpha$ is a root; but by Definition 3.17 this can happen only if $\beta = \alpha$.

If Π' is a second simple set of roots of M in Λ, define $P(\Pi')$ similarly to $P(\Pi)$. If $\Pi \subseteq P(\Pi')$, then, since every element β of Π' may be written as $\beta = \sum_{\alpha \in \Pi} n_\alpha \alpha$ with integers n_α either all positive or all negative, and vice-versa, we must clearly have $\Pi' = \Pi$. We now argue by induction on the number of elements in $P(\Pi)$ but not in $P(\Pi')$. If this is zero, then $\Pi \subseteq P(\Pi')$, so that, as we have seen, $\Pi' = \Pi$. If $P(\Pi)$ is not wholly included in $P(\Pi')$, then neither is Π, and there exists an element $\alpha \in \Pi$,

$\alpha \in P(\Pi')$. By the preceding lemma, $W_\alpha(P(\Pi) \cap P(\Pi')) \subseteq P(\Pi)$. Since W_α is a linear transformation we have $W_\alpha P(\Pi) = P(W_\alpha \Pi)$, and thus $W_\alpha(P(\Pi) \cap P(\Pi')) \subseteq P(\Pi) \cap P(W_\alpha \Pi')$. But $W_\alpha \alpha = -\alpha \in P(\Pi')$ since $\alpha \in P(\Pi')$, and since $W_\alpha^2 = I$ it follows that $\alpha \in P(W_\alpha \Pi')$ also. Hence $P(\Pi) \cap P(W_\alpha \Pi')$ contains at least one more element of Π then does $W_\alpha(P(\Pi) \cap P((W_\alpha \Pi'))$. It follows inductively that there exists a sequence $\alpha_1, \ldots, \alpha_n$ of elements of Π such that $\Pi \subseteq P(W_{\alpha_1} W_{\alpha_2} \cdots W_{\alpha_n} \Pi')$, so that $\Pi = W_{\alpha_1} W_{\alpha_2} \cdots W_{\alpha_n} \Pi'$. Therefore, by Lemma 9 there exists an inner automorphism ϕ mapping M into itself such that the permutation $\alpha \to \alpha'$ of $\{\alpha\}$ defined by $\phi(\mu_\alpha) = \mu_{\alpha'}$ sends Π onto Π', and Theorem 4 is proved.

We record a corollary and prove a lemma for use in a later section.

10. Corollary. The automorphism ϕ of Theorem 4 can be chosen in such a way that $\alpha \to \alpha'$ is a product $W_{\alpha_1} W_{\alpha_2} \cdots W_{\alpha_n}$ of Weyl reflections, with $\alpha_j \in \Pi, j = 1, \ldots, n$.

11. Lemma. Let ϕ be an automorphism which leaves each element of M fixed. Then ϕ is inner; and, in fact, has the form $\phi = \exp(\mathrm{Ad}(\mu))$, where $\mu \in M$.

Proof: Since Λ_α is defined by

$$\Lambda_\alpha = \{\lambda \in \Lambda | [\lambda, \mu] = \alpha(\mu)\, \lambda, \mu \in M\},$$

it is plain that $\phi(\Lambda_\alpha) = \Lambda_\alpha$. Thus, for each $\alpha \in \Pi \cap (-\Pi)$, there exists a constant $c_\alpha \neq 0$ such that $\phi(\lambda_\alpha) = c_\alpha \lambda_\alpha$. Since ϕ must preserve the natural bilinear form of Λ, we have $c(\alpha) = c(-\alpha)^{-1}$. If $\mu \in M$, then $\sigma = \exp(\mathrm{Ad}\,\mu))$ leaves each element of M fixed and maps λ_α into $\exp(\alpha(\mu))\, \lambda_\alpha$. Since by Definition 3.17, Π is a linearly independent set of linear functionals on M, we may choose $\mu \in M$ so that $\alpha(\mu) = -\log c_\alpha, \alpha \in \Pi$. Then plainly $\sigma(\lambda_\alpha) = \phi(\lambda_\alpha)$ for all $\alpha \in \Pi \cap (-\Pi)$, so by Corollary 3.19(c), we have $\sigma(\lambda) = \phi(\lambda)$ for all $\lambda \in \Lambda$. Q.E.D.

6. Representations of Complex Semi-Simple Algebras

1. Definition. Let ϱ be a representation of a Lie algebra Λ on a vector space V. Then

(i) A subspace V_1 of V is called *invariant* (for ϱ) if $\varrho(\lambda)\, V_1 \subseteq V_1$ for each $\lambda \in \Lambda$.

(ii) An invariant subspace V_1 is called *irreducible* if V_1 contains no proper invariant subspace. In the contrary case, V_1 is called *reducible*. If V itself is irreducible, we say that ϱ is irreducible.

Our first main aim is to prove the following important theorem.

2. Theorem. Let ϱ be a representation of the simple Lie algebra Λ in a finite-dimensional vector space V. Then we may write $V = V_1 \oplus \cdots \oplus V_n$, where all the subspaces V_i are invariant and irreducible.

To the proof of Theorem 2, we preface two lemmas.

3. Lemma. Let ϱ be a representation of the simple Lie algebra Λ in a finite dimensional vector space V. Then, if the representation ϱ is not identically equal to zero, the symmetric bilinear form

$$\langle \lambda, \mu \rangle = \mathrm{tr}\,(\varrho(\lambda)\,\varrho(\mu))$$

defined on Λ is non-singular.

4. Lemma. Let ϱ be a representation of the semi-simple Lie algebra Λ in a finite dimensional vector space V. Suppose that W is an irreducible invariant subspace of V, that $\dim(V) = \dim(W) + 1$, and that $\varrho(\lambda)\,V \subseteq W$ for all $\lambda \in \Lambda$. Then there exists a vector $v \in V$, $v \notin W$, such that $\varrho(\lambda)\,v = 0$ for all $\lambda \in \Lambda$.

Proof of Lemma 3. We first consider the case in which Λ is complex.

The reader will verify at once that $\langle \mathrm{Ad}\,(\nu)\,\lambda, \mu \rangle = -\langle \lambda, \mathrm{Ad}\,(\mu)\,\mu \rangle$ for all $\nu, \lambda, \mu \in \Lambda$. Thus the set $\Lambda_0 = \{\lambda | \langle \lambda, \mu \rangle = 0, \mu \in \Lambda\}$ is an ideal in Λ, and therefore, if the bilinear form $\langle \lambda, \mu \rangle$ is not non-singular, $\langle \lambda, \mu \rangle = 0$ for all $\lambda, \mu \in \Lambda$.

In this case, let M be a Cartan subalgebra of Λ. Since Λ is simple, $\Lambda^2 = \Lambda$. Thus, by Lemma 2.11, $\varrho(\mu)$ is nilpotent for all $\mu \in M$. That is, all of V is contained in the weight-space belonging to the weight zero of ϱ. By Lemma 2.10, $\varrho(\lambda)\,V = \{0\}$ if $\lambda \in \Lambda$ is contained in a root space belonging to a non-zero root of M in Λ. But then it is plain that we also have $\varrho(\mu_\alpha) = 0$ for each non-zero root α of M in Λ (cf. Corollary 3.8). But then, using Corollary 3.10 and the fact that Λ is the direct sum of the root spaces of M, it follows that $\varrho(\lambda) = 0$ for all $\lambda \in \Lambda$, proving our lemma in case Λ is complex.

Next consider the case in which Λ is real. We prove just as above that if $\langle \lambda, \mu \rangle$ is not non-singular, $\langle \lambda, \mu \rangle \equiv 0$ for all $\lambda, \mu \in \Lambda$. In this case, let $\hat{\Lambda}$ denote the complexification of Λ. Let $\hat{V}$ be the complexification of V, i.e., the complex vector space of all formal sums $v + iv'$, $v, v' \in V$, addition and multiplication by complex numbers being defined in evident fashion (cf. Definition 1.2). We may define a representation $\hat{\varrho}$ of $\hat{\Lambda}$ on $\hat{V}$ by

$$\hat{\varrho}(\lambda + i\mu)\,(v + iv') = (\varrho(\lambda)\,v - \varrho(\mu)\,v') + i(\varrho(\lambda)\,v' + \varrho(\mu)\,v)$$

for $\lambda + i\mu \in \hat{\Lambda}$, $v + iv' \in \hat{V}$. Then it is plain that

$$\operatorname{tr}(\hat{\varrho}(\lambda + i\mu), \hat{\varrho}(\lambda + i\mu')) = \langle \lambda, \lambda' \rangle - \langle \mu, \mu' \rangle + i(\langle \mu, \lambda' \rangle + \lambda, \mu' \rangle) = 0$$

for $\lambda + i\mu$, $\lambda' + i\mu' \in \Lambda$. Arguing as in the third paragraph of the present proof, we conclude that $\hat{\varrho}(\lambda + i\mu) = 0$ for all $\lambda + i\mu \in \Lambda$; and *a fortiori*, $\varrho(\lambda) = 0$ for all $\lambda \in \Lambda$. Q.E.D.

Proof of Lemma 4. Use Theorem 2.17 to write $\Lambda = \Lambda_1 \oplus \cdots \oplus \Lambda_n$, where each Λ_i is an ideal of Λ, which, regarded as Lie algebra in its own right, is simple. By Lemma 3, we can divide the ideals Λ_i into two classes; those on which the restriction of the symmetric bilinear form $\langle \lambda, \mu \rangle$ is non-singular, and those on which $\varrho(\lambda)$ is identically zero. Suppose the notations are chosen so that the first class consists of $\Lambda_1, \ldots, \Lambda_k$, the second of $\Lambda_{k+1}, \ldots, \Lambda_n$. If $k = 0$ we have nothing to prove, hence we assume $k > 0$. Let $\Lambda' = \Lambda_1 \oplus \cdots \oplus \Lambda_k$, so that Λ' is a subalgebra of Λ. It is plain that the restriction of the bilinear form $\langle \lambda, \mu \rangle$ to Λ' is non-singular. It is equally plain that to establish our lemma, we have only to find a vector $v \in V$, $v \notin W$, such that $\varrho(\lambda') = 0$ for all $\lambda' \in \Lambda'$.

This we do as follows: Let $\lambda_1, \ldots, \lambda_p$ be a basis for Λ'. Let c^{ij} be the inverse of the non-singular matrix $\langle \lambda_i, \lambda_j \rangle$; put $C = \sum_{i,j=1}^{p} c^{ij} \varrho(\lambda_i)\, \varrho(\lambda_j)$, so that C is the so-called *Casimir Operator* of the representation ϱ.

Next note that the Casimir operator C is defined by ϱ, independently of the basis for Λ' which has been chosen. Indeed, if $\mu_1, \ldots, \mu_p$ is a second basis, we may write $\mu_i = \sum_{j=1}^{p} d_i^j \lambda_j$ and $\lambda_i = \sum_{j=1}^{p} d_i^j \mu_j$, where the matrices $\{d_i^j\}$ and $\{\hat{d}_j^i\}$ are each others' inverse. Then $\langle \mu_i, \mu_j \rangle = \sum_{i',j',=1}^{p} d_i^{i'} d_j^{j'} \langle \lambda_{i'}, \lambda_{j'} \rangle$, so that the inverse of the matrix $\langle \mu_i, \mu_j \rangle$ is $\sum_{i',j'=1}^{p} c^{i'j'} \hat{d}_{i'}^{i} \hat{d}_{j'}^{j}$, and the Casimir operator formed with the aid of the basis $\mu_1, \ldots, \mu_p$ is

$$\sum_{i',j',i,j=1}^{p} c^{i'j'} \hat{d}_{i'}^{i} \hat{d}_{j'}^{j} \varrho(\mu_i)\, \varrho(\mu_j) = \sum_{i',j'=1}^{p} c^{i'j'} \varrho(\lambda_{i'})\, \varrho(\lambda_{j'}) = C.$$

For any real number t and any $\lambda' \in \Lambda'$, $\exp(\operatorname{Ad}(t\lambda'))\, \lambda_i = \mu_i$ is a basis of Λ'. By Lemma I.5.3, $\varrho(\mu_i) = \exp(-t\varrho(\lambda'))\, \varrho(\lambda_i) \exp(t\varrho(\lambda'))$. Thus $\langle \mu_i, \mu_j \rangle = \langle \lambda_i, \lambda_j \rangle$, and it follows from what we have just proved that

$$C = \sum_{i,j=1}^{n} c^{ij} \varrho(\mu_i)\, \varrho(\mu_j),$$

so that
$$C = \exp(t\varrho(\lambda'))\left\{\sum_{i,j=1}^{n} c^{ij}\varrho(\lambda_i)\,\varrho(\lambda_j)\right\}\exp(t\varrho(\lambda')).$$

Differentiating with respect to t and putting $t = 0$, we find the identity

$$C\varrho(\lambda') = \varrho(\lambda')\,C, \qquad \lambda' \in \Lambda'. \tag{1}$$

Since we have assumed that $\varrho(\lambda)\,V \subseteq W$ for $\lambda \in \Lambda$, it follows from the definition of C that $CV \subseteq W$. Thus $CW \subseteq W$; if CW were a proper subspace of W, then by (1) it would be invariant. Hence, either $CW = W$ and $C|W$ is a non-singular mapping of W onto itself, or $C = 0$. But $\operatorname{tr}(C) = \sum_{i,j=1}^{p} c^{ij}\operatorname{tr}(\varrho(\lambda_i)\,\varrho(\lambda_j)) = \sum_{i,j=1}^{p} c^{ij}\langle\lambda_i, \lambda_j\rangle = p$, so that $C = 0$ is impossible. Hence $C|W$ is a non-singular mapping of W onto itself. Then plainly $V = \{v|Cv = 0\} \oplus W$. The subspace $\{v|Cv = 0\}$ of V is invariant and *one-dimensional*, so that since $\varrho(\lambda)\,V \subseteq W$ for all $\lambda \in \Lambda$, we must have $\varrho(\lambda)\,\{v|Cv = 0\} = 0$ for $\lambda \in \Lambda$, and thus choosing any non-zero $v\{v|Cv = 0\}$ we establish Lemma 4. Q.E.D.

Now we are able to give the

Proof of Theorem 2. We first show that

(a) If W is an invariant subspace of V such that $1 + \dim(W) = \dim(V)$, and if $\varrho(\lambda)\,V \subseteq W$ for all $\lambda \in \Lambda$, then there exists a vector $v \in V$, $v \notin W$, such that $\varrho(\lambda)\,v = 0$ for all $\lambda \in \Lambda$.

If W is irreducible, this is simply Lemma 4. In the general case, we proceed by induction on $\dim(W)$. If $\dim(W) = 1$, W is irreducible. If $\dim(W) > 1$ and W is not irreducible, it contains proper invariant subspace W'. The elements of the factor space V/W' are the cosets $v + W'$ of W' in V; if we put $\varrho'(\lambda)\,(v + W') = \varrho(\lambda)\,v + W'$, we obtain a representation ϱ' of Λ in V/W'. Let ϕ map each $v \in V$ onto its coset $v + W'$ in V/W'. Then plainly $\phi(W)$ is invariant for ϱ', $\dim(V/W') = \dim(\phi(W)) + 1$, and $\varrho'(\lambda)\,(V/W') \subseteq \phi(W)$ for each $\lambda \in \Lambda$. Thus, by inductive hypothesis, there exists a $v' \in V/W'$ such that $v' \notin \phi(W)$, $\varrho(\lambda)\,v' = 0$, i.e., a $v_0 \in V$ such that $v_0 \notin W$ and $\varrho(v_0) \in W'$. Let $V' = (v_0) + W'$. Then $\varrho(V') \subseteq W'$, and $\dim(V') = \dim(W') + 1$. Hence once more by inductive hypothesis, there exists a $v \in V'$ such that $v \notin W'$ and $\varrho(\lambda)\,v = 0$ for all $\lambda \in \Lambda$. If we write $v = \alpha v_0 + w'$, where $w' \in W' \subseteq W$, it is clear that $\alpha \neq 0$; thus $v \notin W$, and (a) is proved.

Next we prove:

(b) If W is any proper invariant subspace of V, there exists a second invariant subspace W' such that $V = W \oplus W'$.

To do this, we let $\mathscr{T}$ be the set of all linear transformations T in the vector space V such that $TV \subseteq W$, and such that $T|W$ is a constant multiple $\alpha(T)\, I$ of the identity. Let $\mathscr{T}_0$ be the set of all $T \in \mathscr{T}$ such that $\alpha(T) = 0$, i.e., the subspace consisting of all $T \in \mathscr{T}$ such that $T|W = 0$. Plainly, the dimension of $\mathscr{T}$ exceeds that of $\mathscr{T}_0$ by 1. Define a representation σ of Λ on $\mathscr{T}$ by writing $\sigma(\lambda)\, T = \varrho(\lambda)\, T - T\varrho(\lambda)$. Then it is easily verified that $\sigma(\lambda)\, \mathscr{T} \subseteq \mathscr{T}_0$ for $\lambda \in \Lambda$. Thus, by (a), we can find an element $T^+ \in \mathscr{T}$, such that $\alpha(T^+) \neq 0$, and such that $\sigma(\lambda)\, T^+ = 0$ for $\lambda \in \Lambda$, i.e., $\varrho(\lambda)\, T^+ = T^+\varrho(\lambda)$ for all $\lambda \in \Lambda$. We may plainly suppose without loss of generality that $\alpha(T^+) = 1$; and then it is quite clear that $W' = \{v \in V | T^+ v = 0\}$ is invariant, and that $V = W \oplus W'$, proving (b).

It is obvious that Theorem 2 follows by induction on the dimension of V and by use of (b). Q.E.D.

5. Corollary. If W is any proper invariant subspace of V, there exists a second invariant subspace W' such that $V = W \oplus W'$.

6. Corollary. If W is an invariant subspace of V such that $\dim(V) = \dim(W) + 1$, and if $\varrho(\lambda)\, V \subseteq W$ for all $\lambda \in \Lambda$, then there exists a vector $v \in V$, $v \notin W$, such that $\varrho(\lambda)\, v = 0$ for all $\lambda \in \Lambda$.

Remark. Note that Lemma 4 is a slight generalization of Corollary 6.

Our next aim is to analyze the irreducible finite representations of an arbitrary semi-simple complex Lie algebra. *Throughout the remainder of the present section, we let Λ be a semi-simple complex Lie algebra, M a Cartan subalgebra, $\{\alpha\}$ the set of roots of M in Λ, and Π a simple set of roots. For each $0 \neq \alpha \in \{\alpha\}$, we let Λ_α be the one-dimensional root space belonging to the root α; using* Corollary 3.2, *we choose a vector $\lambda_\alpha \in \Lambda_\alpha$ such that $(\lambda_\alpha, \lambda_{-\alpha}) = 1$, where (λ, μ) is the natural bilinear form of Λ. We let $\mu_\alpha = [\lambda_\alpha, \lambda_{-\alpha}]$ be the vector of* Corollary 3.8.

7. Definition. Let ϱ and ϱ_1 be representations of Λ in vector spaces V and V_1 respectively. Then we say that ϱ and ϱ_1 are *equivalent* if there exists a one-to-one linear transformation T of V onto V_1, such that $T\varrho(\lambda)\, T^{-1} = \varrho_1(\lambda)$, $\lambda \in \Lambda$.

Of course, we aim to classify irreducible representations up to equivalence. The following definition introduces the objects upon which our classification will be based.

8. Definition. Let ϱ be a representation of Λ in a vector space V. Then weights of the representation $\mu \to \varrho(\mu)$ of M will be called (for brevity) the *weights of* ϱ. If ω is a weight of ϱ, and if $v \in V$ is a weight vector belonging to the weight ω such that $\varrho(\mu)\, v = \omega(\mu)\, v$ for $\mu \in M$, then v will be called a *minimal weight vector* if $\varrho(\lambda_{-\alpha})\, v = 0$ for all $\alpha \in \Pi$; in this case, ω will be called a *minimal weight*.

The next lemma, which shows that an irreducible representation is characterized by its minimal weight, is proved by a simple adaptation of the line of proof of the uniqueness lemma 3.20.

9. Lemma. Let ϱ_1 and ϱ_2 be two irreducible representations of Λ, in vector spaces V_1 and V_2 respectively. Suppose that both representations have a weight ω, and that ω is a minimal weight of both representations. Then ϱ_1 and ϱ_2 are equivalent.

Proof: Let $v_1 \in V_1$ and $v_2 \in V_2$ be minimal weight vectors belonging to the weight ω of ϱ_1 and ϱ_2 respectively. Let $V = V_1 \oplus V_2$, and define the representation ϱ of Λ in V by $\varrho(\lambda)\, [\hat{v}_1, \hat{v}_2] = [\varrho_1(\lambda)\, \hat{v}_1, \varrho_2(\lambda)\, \hat{v}_2]$, $\hat{v}_1 \in V_1$, $\hat{v}_2 \in V_2$. Then clearly $v = [v_1, v_2]$ is a minimal weight vector of ϱ belonging to the minimal weight ω. Let W be the least invariant subspace of V containing v. Since V_1 (resp. V_2) is irreducible, it is plain that the natural projection of W on V_1 (resp. V_2) is all of V_1 (resp. V_2). We shall show that $W \subseteq V_1 \oplus V_2$ is the graph of a one-to-one mapping of V_1 onto V_2. To do this, we have only to show that W contains no vector of the form $[\hat{v}_1, 0]$ or $[0, \hat{v}_2]$, where $0 \neq \hat{v}_1 \in V_1$, $0 \neq \hat{v}_2 \in V_2$. But, if W contained a vector $[\hat{v}_1, 0]$, $\hat{v}_1 \neq 0$, then since ϱ_1 is irreducible W would contain all of V_1, hence contain $[v_1, 0]$, hence contain $[v_1, v_2] - [v_1, 0] = [0, v_2]$, hence contain all of V_2, and hence contain all of V. This makes it plain that to show that W is the graph of a one-to-one mapping of V_1 onto V_2, we have only to show that W is a proper subspace of V. We will in fact prove that the only vectors in W belonging to the minimal weight ω are scalar multiples of v.

This we do as follows: since W is the least invariant subspace of V which contains v, and since by Corollary 3.19 Λ is generated by the elements λ_α, $\alpha \in \Pi \cup (-\Pi)$, W is evidently spanned by the vectors $\varrho(\lambda_{\alpha_n}) \ldots \varrho(\lambda_{\alpha_1})\, v$, where $\alpha_j \in \Pi \cup (-\Pi)$, $1 \leqslant j \leqslant n$. By Lemma 2.10, such a vector belongs to the weight $\omega + \sum_{i=1}^{n} \alpha_i$ of ϱ. Therefore, if a finite linear combination

$$\sum_{n} \sum_{\alpha_1, \ldots, \alpha_n \in \pi \cup (-\pi)} c(\alpha_n, \ldots, \alpha_1)\, \varrho(\lambda_{\alpha_n}) \ldots \varrho(\lambda_{\alpha_1})\, v$$

is a weight vector belonging to weight ω, it must equal the partial sum

$$\sum_{n} \sum_{\substack{\alpha_1, \ldots, \alpha \in \Pi \cup (-\Pi) \\ \alpha_1 + \cdots + \alpha_n = 0}} c(\alpha_n, \ldots, \alpha_1)\, \varrho(\lambda_{\alpha_n}) \ldots \varrho(\lambda_{\alpha_1})\, v.$$

Therefore, we have only to show that if $\sum_{i=1}^{n} \alpha_i = 0$, then $\varrho(\lambda_{\alpha_n}) \cdots \varrho(\lambda_{\alpha_1})\, v$ is a multiple of v.

To this end, we proceed inductively as follows. Define the *rank* of the sequence $\alpha_1, \ldots, \alpha_n$ as $\sum_{\alpha_i \in -\Pi} i$. By Definition 3.17, Π is a linearly independent set of linear functionals on M. Thus, if $\sum_{i=1}^{n} \alpha_i = 0$, as many elements of Π as of $(-\Pi)$ must appear in the sequence $\alpha_1, \ldots, \alpha_n$. Therefore, if $\alpha_1, \ldots, \alpha_n$ has lowest possible rank, then $\alpha_1 \in (-\Pi)$, $\varrho(\lambda_{\alpha_1})\, v = 0$, and our assertion is verified. Now argue by induction on the rank. If the rank is more than minimal, there will exist a least one pair α_i, α_{i+1} with $\alpha_{i+1} \in (-\Pi)$ and $\alpha_i \in \Pi$. If $\alpha_{i+1} \neq -\alpha_i$, then $[\lambda_{\alpha_i}, \lambda_{\alpha_{i+1}}] = 0$ by Lemma 3.11 and Definition 3.17, so that

$$\varrho(\lambda_{\alpha_n}) \cdots \varrho(\lambda_{\alpha_{i+1}})\, \varrho(\lambda_{\alpha_i}) \cdots \varrho(\lambda_1)\, v = \varrho(\lambda_{\alpha_n}) \cdots \varrho(\lambda_{\alpha_i})\, \varrho(\lambda_{\alpha_{i+1}}) \cdots \varrho(\lambda_{\alpha_i})\, v, \quad (2)$$

which by inductive hypothesis gives the desired conclusion. If $\alpha_{i+1} = -\alpha_i$, then by Corollary 3.8 we have $[\lambda_{\alpha_i}, \lambda_{\alpha_{i+1}}] = \mu_{\alpha_i}$, so that

$$\begin{aligned}
\varrho(\lambda_{\alpha_n}) \cdots \varrho(\lambda_{\alpha_{i+1}})\, \varrho(\lambda_{\alpha_i}) \cdots \varrho(\lambda_1)\, v &= \varrho(\lambda_{\alpha_n}) \cdots \varrho(\lambda_{\alpha_{i+1}})\, \varrho(\lambda_{\alpha}) \cdots \varrho(\lambda_1)\, v \\
&\quad - \varrho(\lambda_{\alpha_n}) \cdots \varrho(\lambda_{\alpha_{i+2}})\, \varrho(\mu_{\alpha_i})\, \varrho(\lambda_{\alpha_{i+1}}) \cdots \varrho(\lambda_1)\, v \\
&= \varrho(\lambda_{\alpha_n}) \cdots \varrho(\lambda_{\alpha_i}) \cdots \varrho(\lambda_{\alpha_{i-1}}) \cdots \varrho(\lambda_1)\, v \\
&\quad - [\alpha_1(\mu_{\alpha_i}) + \cdots + \alpha_{i-1}(\mu_{\alpha_i}) \\
&\quad + \omega(\mu_{\alpha_i})]\, \varrho(\lambda_{\alpha_n}) \cdots \varrho(\lambda_{\alpha_{i+2}}) \varrho(\lambda_{\alpha_{i-1}}) \cdots \varrho(\lambda_1)\, v, \qquad (3)
\end{aligned}$$

which give the desired conclusion by inductive hypothesis once more. This completes our induction, and with it our proof that W is the graph of a one-to-one mapping of V_1 on V_2.

Let the mapping of V_1 on V_2 whose graph is W be T. Since W is invariant under ϱ, we have $T\varrho_1(\lambda)\, v_1 = \varrho_2(\lambda)\, Tv_1$ for each $v_1 \in V_1$ and $\lambda \in \Lambda$. Thus $T\varrho_1(\lambda)\, T^{-1} = \varrho_2(\lambda)$, and ϱ_1 is equivalent to ϱ_2. Q.E.D.

The reader will have no difficulty in adapting the above line of argument, based on induction on the rank of a sequence $\alpha_1, \ldots, \alpha_n$ of elements of $\Pi \cup (-\Pi)$, as defined in the preceding proof, to obtain the additional conclusion stated in the following corollary.

10. Corollary. Let ϱ be an irreducible representation of Λ in a space V, ω is a minimal weight for ϱ, and v is a weight vector belonging to this minimal weight

(a) The vector v is unique up to a scalar multiple.
(b) Let $\alpha_1, \ldots, \alpha_n \in \Pi \cup (-\Pi)$ and $\alpha_1 + \cdots + \alpha_n = \sum_{\alpha \in \Pi} k_\alpha \alpha$. Then if any of the integers k_α is negative, $\varrho(\lambda_{\alpha_n}) \cdots \varrho(\lambda_{\alpha_1})\, v = 0$.
(c) Every weight of ϱ is of the form $\omega + \sum_{\alpha \in \Pi} k_\alpha \alpha$, where each of the integers k_α is non-negative.
(d) V is spanned by vectors of the form $x = \varrho(\lambda_{\alpha_n}) \cdots \varrho(\lambda_{\alpha_1})\, v$, where $\alpha_1, \ldots, \alpha_n \in \Pi$, each of which is a weight vector belonging to the weight $\omega + \sum_{i=1}^{n} \alpha_i$ of ϱ.
(e) If v_1 is a weight vector belonging to the weight ω_1 of Λ, then $\varrho(\mu)\, v_1 = \omega_1(\mu)\, v_1$ for each $\mu \in M$.
(f) The weight spaces of ϱ are finite dimensional.

An immediate consequence of Corollary 10(c) is

11. Corollary. The minimal weight ω of an irreducible representation ϱ of Λ admitting a minimal weight is unique.

Note that in proving Lemma 9 and Corollaries 10 and 11 we need never require that V is finite dimensional. Thus, all three of these results are available even if V is infinite dimensional. This observation will be of use to us very shortly.

If we assume that V is finite dimensional, we can say a little more.

12. Lemma. Let ϱ be an irreducible representation of Λ in a finite dimensional space V. Then ϱ admits a unique minimal weight ω and we have

$$\omega(\mu_\alpha) = -\tfrac{1}{2} m_\alpha \alpha\,(\mu_\alpha), \qquad \alpha \in \Pi$$

where m_α is a non-negative integer.

Proof. Since V is finite dimensional, it follows by that V decomposes into the direct sum of its weight spaces V_ω, which are finite in number. By Lemma 2.10, $\varrho(\lambda_{-\alpha})\, v \in V_{\omega-\alpha}$ if $v \in V_\omega$. Thus, if there were no minimal weight, there would be weights of the form $\omega - \sum_{\alpha \in \Pi} n_\alpha \alpha$, with $n_\alpha \geqslant 0$ and $\sum_{\alpha \in \Pi} n_\alpha$ arbitrarily large. This is impossible, since by Definition 3.17 Π is a linearly independent set of linear functionals on M. This proves the first assertion of our lemma.

To prove the second assertion, we let v be a vector belonging to the minimal weight ω. Let $\alpha \in \{\alpha\}$, $\alpha \neq 0$. Then, from what we have said, there must exist some least finite integer m_α such that $\varrho(\lambda_\alpha)^{m_\alpha+1}\, v = 0$. The subspace V_0 of V generated by the vectors $v, \varrho(\lambda_\alpha)\, v, \ldots\, \varrho(\lambda_\alpha)^{m_\alpha} v$ is easily seen to be invariant under λ_α and $\lambda_{-\alpha}$. Thus,

$$0 = \operatorname{tr}\left((\varrho(\lambda_\alpha)|V_0)\,(\varrho(\lambda_{-\alpha})|V_0) - (\varrho(\lambda_{-\alpha})|V_0)\,(\varrho(\lambda_\alpha)|V_0)\right)$$

$$= \operatorname{tr}(\varrho[\lambda_\alpha, \lambda_{-\alpha}]|V_0) = \operatorname{tr}(\varrho(\mu_\alpha)|V_0) = \sum_{k=0}^{m_\alpha} \omega(\mu_\alpha) + k\alpha(\mu_\alpha),$$

Q.E.D.

The following theorem is now decisive.

13. Theorem. Suppose that for each finite dimensional irreducible representation ϱ of Λ we let ω_ϱ be its unique minimal weight and write

$$\omega_\varrho(\mu_\alpha) = -\tfrac{1}{2} m_\alpha\, \alpha(\mu_\alpha), \qquad \alpha \in \Pi. \tag{4}$$

Then (4) establishes a $1 - 1$ correspondence between the family of all indexed sets of non-negative integers m_α, $\alpha \in \Pi$, and the family of all classes of equivalent irreducible finite-dimensional representations of Λ.

That the correspondence established by (4) is one-to-one is the burden of Lemma 9 and Corollary 11; thus only the existence of a representation corresponding to each indexed set of non-negative integers must be proved. For this very concrete existence assertion, we will find it convenient to give a highly transfinite proof! We divide the proof into a series of lemmas.

14. Lemma. Let ϱ be a representation of Λ in a space V, and suppose that ϱ admits a minimal weight ω. Then there exists an irreducible representation ϱ' of Λ in a space V' which admits the minimal weight ω.

Proof. Let $0 \neq v$ be a minimal weight vector belonging to the weight ω, so that $\varrho(\mu)\, v = \omega(\mu)\, v$ for all $\mu \in M$. Let V_0 be the smallest invariant subspace of V containing v, and let $\varrho_0(\lambda) = \varrho(\lambda)|V_0$. It is plain that $v \in V_0$ is a minimal weight vector belonging to the weight ω of ϱ_0. Clearly, an invariant subspace of V_0 is proper if and only if it fails to contain the vector v. Thus by Zorn's lemma, V_0 contains a maximal proper invariant subspace V_1. Let V' be the factor space V_0/V_1, i.e., the set of all cosets $u + V_1, u \in V_0$, and let ϕ be the natural mapping $u \to u + V_1$ of V_0 onto V'. Define a representation ϱ' of Λ on V' by $\varrho'(\lambda)\,(u + V_1) = \varrho(\lambda)\, u + V_1$, $u \in V_0$, $\lambda \in \Lambda$. Then it is easily verified that $V_2 \leftrightarrow \phi(V_2) = V_2'$ establishes a one-to-one correspondence between invariant subspaces V_2' of V' and invariant subspaces V_2 of V_0 containing V_1. But then, V_1 is a maximal invariant proper subspace of V_0, V' has no proper invariant spaces, and is therefore irreducible. Q.E.D.

Note that in Lemma 14 we neither assume nor assert finite-dimensionality. But if V is assumed finite dimensional, it is clear that V' is finite dimensional also.

Using Lemma 14 and the preceding remark, we may show that a great many indexed sets of non-negative integers do correspond to irreducible finite dimensional representations, in the sense of (4) of Theorem 13. We proceed as follows: Let ϱ and ϱ' be finite dimensional irreducible representations of $\varLambda$ in spaces V and V' respectively, and let ω and ω' be their respective minimal weights. Let W' be the space of bilinear functionals $\varPhi(v, v')$, $v \in V$, $v' \in V'$, and let W be the space of linear functionals on W', the so called *Kronecker product* of V and V'. There is a natural bilinear map $[v, v'] \to v \otimes v'$ of $V \times V'$ into W, defined by $(v \otimes v')(\varPhi) = \varPhi(v, v')$. Let ϱ' be the representation of $\varLambda$ in W' defined by $(\varrho'(\lambda)\varPhi)(v,v') = -\varPhi(\varrho(\lambda)v, v') - \varPhi(v, \varrho'(\lambda)v')$, and let $\hat{\varrho}$ be the representation of $\varLambda$ in W defined by $(\hat{\varrho}(\lambda)\varPhi')(\varPhi) = -\varPhi'(\varrho'(\lambda)\varPhi)$, $\varPhi' \in W$, $\varPhi \in W'$, $\lambda \in \varLambda$. Then it is immediately verified that

$$\hat{\varrho}(\lambda)(v \otimes v') = (\varrho(\lambda)v \otimes v') + v \otimes \varrho'(\lambda)v'. \tag{5}$$

Suppose now that we choose v and v' to be minimal weight vectors of ϱ and ϱ' belonging to ω and ω' respectively. Then, by (5),

$$\hat{\varrho}(\lambda_{-\alpha})(v \otimes v') = 0 \quad \alpha \in \varPi;$$

and

$$\varrho(\mu)(v \otimes v') = (\omega(\mu) + \omega'(\mu))\, v \otimes v', \ \mu \in M.$$

Thus, using the preceding lemma and remark, we obtain the following statement.

15. Lemma. If $\varLambda$ has irreducible finite dimensional representations admitting the minimal weights ω and ω' respectively, it has an irreducible finite dimensional representation admitting the minimal weight $\omega + \omega'$.

By generalizing the above construction, we can obtain still more minimal weights. Let ϱ be a representation of $\varLambda$ in a finite-dimensional space V. For each weight ω' of ϱ, let $V_{\omega'}$ denote the weight space belonging to ω'. Let n be any integer, and let X'' be the space of all functionals $\varPhi(v_1, \ldots, v_n)$ of n variables $v_j \in V$, separately linear in each variable. Let X' be the space of linear functionals on X''. There is a natural multilinear map $[v_1, \ldots, v_n] \to v_1 \otimes \cdots \otimes v_n$ of $V \times \cdots \times V$ into X', defined by

$$(v_1 \otimes \cdots \otimes v_n)\,\Phi = \Phi(v_1, \ldots, v_n).$$

The reader will easily verify that if $u_1, \ldots, u_k$ is a basis for V, then $\{u_{i_1} \otimes \cdots \otimes u_{i_n}\}$, where $1 \leqslant i_j \leqslant k, j = 1, \ldots, n$ is a basis for X'. We let $\check{\varrho}$ be the representation of Λ in X'' defined by

$$(\check{\varrho}(\lambda)\,\Phi)\,(v_1, \ldots, v_n) = -\sum_{j=1}^{n} \Phi(v_1, \ldots, \varrho(\lambda)\,v_j, \ldots, v_n), \tag{6}$$

and let ϱ be the representation of Λ in X' defined by $(\hat{\varrho}(\lambda)\,\Phi')\,(\Phi) = -\Phi'(\check{\varrho}(\lambda)\,\Phi)$, $\Phi' \in X'$, $\Phi \in X''$, $\lambda \in \Lambda$. Then it is immediately verified that

$$\hat{\varrho}(\lambda)\,(v_1 \otimes \cdots \otimes v_n) = \sum_{j=0}^{n} v_1 \otimes \cdots \otimes \varrho(\lambda)\,v_j \otimes \cdots \otimes v_n. \tag{7}$$

Let $\mathscr{P}$ denote the group of all permutations of $1, \ldots, n$, i.e., the group of all one-to-one mappings of this set onto itself. If $\pi \in \mathscr{P}$, we may define a linear mapping $\check{\pi}$ on X'' by $(\check{\pi}\Phi)\,(v_1, \ldots, v_n) = \Phi(v_{\pi^{-1}(1)}, \ldots, v_{\pi^{-1}(n)})$, and a linear mapping $\hat{\pi}$ in X' by $(\hat{\pi}\Phi')\,(\Phi) = \Phi'(\check{\pi}^{-1}\Phi)$, $\Phi \in X$, $\Phi \in X'$. Then it is easily verified that

$$\hat{\pi}(v_1 \otimes \cdots \otimes v_n) = v_{\pi(1)} \otimes \cdots \otimes v_{\pi(n)}. \tag{8}$$

It follows by (7) that $\hat{\pi}\hat{\varrho}(\lambda) = \hat{\varrho}(\lambda)\,\hat{\pi}$ if $\pi \in \mathscr{P}$ and $\lambda \in \Lambda$. Thus, if sgn (π) denotes the sign of the permutation π, the subspace X_0 of X defined by

$$X_0' = \{x \in X' | \hat{\pi}x = \operatorname{sgn}(\pi)\,x, \quad \pi \in \mathscr{P}\}$$

is invariant for $\hat{\varrho}$. If $u_1, \ldots, u_k$ is a basis for V, then since $\{u_{i_1} \otimes \cdots \otimes u_{i_n}\}$, where $1 \leqslant i \leqslant k, j = 1, \ldots, n$ is a basis for X', it easily verified that

$$\left\{\frac{1}{n!} \sum_{\pi \in \mathscr{P}} u_{i_{\pi(1)}} \otimes \cdots \otimes u_{i_{\pi(n)}}\right\}, \tag{9}$$

where $1 \leqslant i_1 < i_2 < i_3 < \cdots < i_k \leqslant k$, is a basis for X_0'. Using (7) and Corollary 10(d) and (e), we see that the set of weights $\omega^{(0)}$ of the representation $\lambda \to \varrho_0(\lambda) = \hat{\varrho}(\lambda)|X_0'$ is the collection of all sums $\omega^{(0)} = \omega_1 + \omega_2 + \cdots + \omega_n$, where $\omega_1, \ldots, \omega_n$ are weights of ϱ, no weight ω_j in this list being repeated more often than dim (V_{ω_j}).

Suppose now that ϱ is irreducible, and that its minimal weight is ω, so that by Corollary 10(c) each of its weights is of the form $\omega + \sum_{\alpha \in \pi} k_\alpha \alpha$, where the integers k_α are all non-negative. Then we may introduce a "lexicographic" ordering into the collection of weights of ϱ, as follows: Enumerate the roots α in any order, and then write $\omega + \sum_{\alpha\in\Pi} k_\alpha\alpha > \omega + \sum_{\alpha\in\Pi} k'_\alpha\alpha$ if at the first $\alpha \in \Pi$ for which $k_\alpha - k'_\alpha$ is different from zero we have $k_\alpha - k'_\alpha > 0$. The weights of ϱ_0 have the form $n\omega + \sum_{\alpha\in\Pi} k_\alpha\alpha$, and may be ordered by the same lexicographic principle. The least weight of ϱ_0 in lexicographic order is then $\omega_1 + \cdots + \omega_n$, where $\omega_1, \ldots, \omega_n$ are the first n weights of ϱ in the lexicographically ordered list of weights in which each weight ω' is repeated exactly dim $(V_{\omega'})$ times.

Since $\omega_1 + \cdots + \omega_n$ is the least weight of ϱ_0 in lexicographic order, it follows by Lemma 2.10 that it is a minimal weight of ϱ_0. Thus, using Lemma 14 and the remark which follows it, we see that for each n there exists a finite dimensional irreducible representation whose minimal weight is $\omega_1 + \cdots + \omega_n$, provided that n does not exceed the length of the entire lexicographically ordered list. Applying this to the adjoint representation of Λ, whose weights are the roots of M in Λ, and using Lemma 3.7, we obtain the following lemma.

16. Lemma. Let the set Π of simple roots be enumerated in some order, and then order the non-zero roots $\alpha \in \{\alpha\}$ by writing $\beta_1 > \beta_2$ if $\beta_1 = \sum_{\alpha\in\Pi} n_\alpha\alpha, \beta_2 = \sum_{\alpha\in\Pi} n'_\alpha\alpha$, and if at the first $\alpha \in \Pi$ for which $n_\alpha - n'_\alpha \neq 0$ we have $n_\alpha > n'_\alpha$. Then, for each integer n not exceeding the number of non-zero roots $\alpha \in \{\alpha\}$, the sum $\alpha_1 + \cdots + \alpha_n$, when $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the first n non-zero roots in lexicographic order, is the minimal weight of some irreducible finite-dimensional representation of Λ.

Combining Lemmas 15 and 16, we obtain the following Corollary.

17. Corollary. Suppose that the non-zero roots $\alpha \in \{\alpha\}$ are ordered as in the preceding lemma. If n_α is an integer for each $0 \neq \alpha \in \{\alpha\}$, and if $n_\alpha \geqslant n_\beta$ whenever $\alpha > \beta, \alpha, \beta \in \{\alpha\}, \alpha \neq 0 \neq \beta$, then the sum $\sum_{\substack{\alpha\in\{\alpha\}\\ \alpha\neq 0}} n_\alpha\alpha$ is the minimal weight of an irreducible finite-dimensional representation of Λ.

Remark. The representation whose existence is asserted in the preceding corollary may in fact be obtained from the adjoint representation $\lambda \to \mathrm{Ad}\,(\lambda)$ by iterated application of the constructions described in the paragraphs immediately preceding and immediately following the statement of Lemma 15.

Now we prepare to give our argument a transcendental turn. Let m

be the dimension of M. For notational simplicity, we will write an m-triple $[z_1, \ldots, z_m]$ of complex numbers as Z, correspondingly write $[az_1 + bz'_1, \ldots, az_m + bz'_m]$ as $aZ + bZ'$, and write a polynomial $P(z_1, \ldots, z_n)$ as $P(Z)$. If $\Pi = \{\alpha_1, \ldots, \alpha_m\}$ is the ordering of Π to which reference is made in Lemma 16, and ω is a linear functional on M, we may also write $[\omega(\mu_{\alpha_1}), \ldots, \omega(\mu_{\alpha_n})]$ as $\omega(\mu_\alpha)$. We use these simplified notations in stating the following lemma, which recasts Corollary 17 in a more useful algebraic form.

18. Lemma. Let $P(z_1, \ldots, z_m) = P(Z)$ be a polynomial. Then if $P(\omega(\mu_\alpha)) = 0$ whenever ω is the minimal root of an irreducible representation of Λ, $P(Z) = 0$ identically.

Proof. Let $\beta_1, \ldots, \beta_n$ be the non-zero roots of M in Λ, enumerated in the lexicographic order of Corollary 17. Then, by Corollary 17, the polynomial $Q(\zeta_1, \ldots, \zeta_n) = P(\zeta_1\beta_1(\mu_\alpha) + \cdots + \zeta_n\beta_n(\mu_\alpha))$ vanishes whenever $\zeta_1, \ldots, \zeta_n$ is a decreasing sequence of non-negative integers. Hence it vanishes identically in $\zeta_1, \ldots, \zeta_n$. Since the m vectors μ_α, $\alpha \in \Pi$ are a basis for M by Definition 3.17 and Corollary 3.8, and since the m roots α, $\alpha \in \Pi$ are linearly independent linear functionals on M by Definition 3.17, the present lemma follows. Q.E.D.

We now form a highly infinite-dimensional representation of Λ, as follows: Choose one representation ϱ from each class of equivalent irreducible finite-dimensional representations, thereby getting a class of representations R (countable, by Lemma 12 and Lemma 9). Each representation $\varrho \in R$ acts in a space V_ϱ; let V^* be the space of all functions f which map $R \to \bigcup_{\varrho \in R} V_\varrho$, and are such that $f(\varrho) \in V_\varrho$, $\varrho \in R$. Make V^* into a vector space by setting $(af + bg)(\varrho) = af(\varrho) + bg(\varrho)$. Define a representation ϱ^* of Λ in V^* by putting $(\varrho^*(\lambda)f)(\varrho) = \varrho(\lambda)(f(\varrho))$, $\lambda \in \Lambda$. For each $\varrho \in R$, let ω_ϱ be the minimal weight of ϱ, and let v_ϱ be a minimal weight vector belonging to the weight ω_ϱ. Define $v^* \in V^*$ by $v^*(\varrho) = v_\varrho$. In terms of these notations, we state the following lemma.

19. Lemma. Suppose that for each $\alpha \in \Pi$ a complex constant c_α is given. Then the vector v^* does not belong to the least invariant subspace V_0^* of V^* generated by the m vectors $(\varrho^*(\mu_\alpha) - c_\alpha)\, v^*$, $\alpha \in \Pi$.

20. Corollary. Suppose that for each $\alpha \in \Pi$ a complex constant c_α is given. Then there exists a unique, though possibly infinite-dimensional, irreducible representation of Λ, possessing as minimal weight the linear functional ω defined by $\omega(\mu_\alpha) = c_\alpha$, $\alpha \in \Pi$.

To lend interest to the proof of the lemma, we first prove the corollary.

Proof of Corollary 20. Let v^*, V^* and V_0^* be as in Lemma 19, and let ϱ^* be as in the paragraph preceding Lemma 19. Define a representation ϱ_1 of Λ in the factor space $V_1 = V^*/V_0^*$, whose elements are the cosets $u^* + V_0^*$ $u^* \in V^*$, by putting $\varrho_1(\lambda)\,(u^* + V_0^*) = \varrho^*(\lambda)\,u^* + V_0^*$, $\lambda \in \Lambda$. It is clear from the definition of v^* that $\varrho^*(\lambda_{-\alpha})\,v^* = 0$, $\alpha \in \Pi$. By Lemma 19, the coset $v_1 = v^* + V_0^*$ is non-zero. Plainly $\varrho_1(\lambda_{-\alpha})\,v_1 = 0$, $\alpha \in \Pi$, and $\varrho_1(\mu_\alpha)\,v_1 = c_\alpha v_1$, $\alpha \in \Pi$. Thus v_1 is a minimal weight vector belonging to the minimal weight ω defined by $\omega(\mu_\alpha) = c_\alpha$. Corollary 20 now follows at once by Lemma 14. Q.E.D.

Proof of Lemma 19. It follows easily by Corollary 3.19 that V_0^* is spanned by the collection of vectors of the form

$$\varrho^*(\lambda_{\alpha_k}) \cdots \varrho^*(\lambda_{\alpha_1})\,(\varrho^*(\mu_{\alpha_{k+1}}) - c_{\alpha_{k+1}})\,v^*, \tag{10}$$

where $\alpha_1, \ldots, \alpha_{k+1} \in \Pi \cup (-\Pi)$. For each $\alpha \in \{\alpha\}$, let $V_{(\alpha)}^*$ denote the subspace of elements $f \in V^*$ such that $f(\varrho) \in V_{\omega_\varrho+\alpha}$ for all $\varrho \in R$. If $f \in \sum_{\beta\neq\alpha} V_{(\beta)}^*$, then $f(\varrho) \in \sum_{\beta\neq\alpha} V_{\omega_\varrho+\beta}$. It is clear from this and from the linear independence of the spaces $V_{\omega_\varrho+\beta}$, cf. Theorem 1.10, that $V_{(\alpha)}^*$ and $\sum_{\beta\neq\alpha} V_{(\beta)}^*$ have no non-zero vector in common. By Lemma 2.10, and from the definition of ϱ^*, $\varrho^*(\lambda_{\alpha_k}) \cdots \varrho^*(\lambda_{\alpha_1})(\varrho^*(\beta_{\alpha_{k+1}}) - c_{\alpha_{k+1}})\,v^* \in V_{(\alpha)}^*$, where $\alpha = \alpha_1 + \cdots + \alpha_k$. Thus, if v^* is a linear combination of elements of the form (10), it is a linear combination of elements of the form (10) with $\sum_{j=1}^{k} \alpha_j = 0$.

We now prove the following:

Auxiliary Statement: If in (10) we have $\alpha_1 + \cdots + \alpha_k = 0$, then the the element (10) is a linear combination of elements of the form $\varrho^*(\mu_1) \cdots \varrho^*(\mu_n)\,(\varrho^*(\mu_\alpha) - c_\alpha)\,v$, where $\mu_1, \ldots, \mu_n \in M$, and $\alpha \in \Pi$.

If $k = 2$ so that $\alpha_1 = -\alpha_2$, this follows from Lemma 3.11 and Corollary 3.9. In the general case, we may proceed inductively, as follows. Define the *rank* of the sequence $\alpha_1, \ldots, \alpha_k$ as $\sum_{\alpha_i \in (-\Pi)} i$. By Definition 3.17, Π is a linearly independent set of functionals on M. Thus, if $\sum_{i=1}^{k} \alpha_i = 0$, as many elements of Π as of $(-\Pi)$, must appear in the sequence $\alpha_1, \ldots, \alpha_n$. If $\alpha_1, \ldots, \alpha_k$ is of lowest possible rank, then we must have $\alpha_1 \in -\Pi$ and thus the element (10) equals

$$\left(\varrho^*(\mu_{\alpha_{k+1}}) - c_{\alpha_{k+1}} - \sum_{j=1}^{k} \alpha_j(\mu_{\alpha_{k+1}})\right) \varrho^*(\lambda_{\alpha_k}) \cdots \varrho^*(\lambda_{\alpha_1})\,v^* = 0,$$

and our assertion is verified. If the rank is more than minimal, there will exist at least one pair α_i, α_{i+1} with $\alpha_{i+1} \in (-\Pi)$ and $\alpha_i \in \Pi$. Then an evident induction on the rank using Lemma 3.11, Corollaries 3.8 and 3.9, and Definition 3.17, (cf. (2) and (3) of the proof of Lemma 9) will establish our auxiliary statement.

Using the auxiliary statement and the fact that $\{\mu_\alpha, \alpha \in \Pi\}$ is a basis for M (cf. Definition 3.17 and Corollary 3.8) we see that to prove the present lemma it is sufficient to show that no equation (11)

$$v^* + \sum_{n=1}^{p} \sum_{i_1,\dots,i_{n+1}=1}^{m} d(i_1,\dots,i_{n+1})\varrho^*(\mu_{i_1})\dots\varrho^*(\mu_{i_n})(\varrho^*(\mu_{i_{n+1}}) - c_{i_{n+1}}) = 0 \tag{11}$$

can hold; here $\mu_1, \dots, \mu_n$ is an enumeration of the vectors μ_α, $\alpha \in \Pi$, and we have written c_i for c_α if $\mu_\alpha = \mu_i$ to spare the sensibilities of the printer; the $d(i_1, \dots, i_n)$ are arbitrary complex coefficients. But, if (11) holds, then it follows from the definition v^* and ϱ^* that

$$1 + \sum_{n=1}^{p} \sum_{i_1,\dots,i_{n+1}}^{m} d(i_1, \dots, i_{n+1})\,\omega(\mu_{i_1}) \cdots \omega(\mu_{i_n})\,(\omega(\mu_{i_{n+1}}) - c_{i_{n+1}}) = 0 \tag{12}$$

whenever ω *is the minimal root of an irreducible representation.* By Lemma 18, this is unless the coefficients d are zero impossible; and our lemma follows. Q.E.D.

Corollary 20 gives the existence and uniqueness (but not the finite-dimensionality) of the representations whose existence and finite-dimensionality is asserted in Theorem 13. Thus all that remains is to prove that if the minimal weight ω of an irreducible representation ϱ satisfies $\omega(\mu_\alpha) = -\frac{1}{2}k_\alpha\alpha(\mu_\alpha)$, $\alpha \in \Pi$, each k_α being a non-negative integer, then ϱ is necessarily finite dimensional.

We verify this assertion first in case our Lie algebra Λ is the three-dimensional algebra Λ_0 of type A_2 of all 2×2 matrices of trace zero. The natural bilinear form on Λ_0 is $(\lambda_1, \lambda_2) = 4\,\mathrm{tr}\,(\lambda_1\lambda_2)$, and using this, the natural representation of Λ_0 by the transformations $\varrho(\lambda): v \to \lambda v$ on 2-dimensional space is seen to have a minimal weight ω satisfying $\omega(\mu_\alpha) = -\frac{1}{2}\alpha(\mu_\alpha)$; here, α is the unique root of Λ_0. Then, by Lemma 15, if $n \geqslant 1$ Λ_0 has a *finite dimensional* representation whose minimal weight ω' is $n\omega$; and our assertion follows by the uniqueness stated in Corollary 20. Note that Λ_0 has a basis consisting of three vectors $\lambda_\pm$, μ, such that $[\mu, \lambda_\pm] = \pm\lambda_\pm$, $[\lambda_+, \lambda_-] = \mu$. The one-dimensional space M spanned by μ is then a Cartan subalgebra, the two non-zero roots of M in Λ_0 are defined by $\alpha_\pm(\mu) = \pm 1$; $\mathrm{tr}\,(\mathrm{Ad}\,\lambda_+\ \mathrm{Ad}\,\lambda_-) = 2$, and $\mathrm{tr}\,(\mathrm{Ad}\,\mu\ \mathrm{Ad}\,\mu)$

$= 2$. Thus, any irreducible representation of Λ_0 which has a minimal weight vector belonging to a weight ω for which $\omega(\mu) = -\frac{n}{2}$, where n is a non-negative integer, is finite-dimensional.

Next we use the fact just proved to establish the following lemma.

21. Lemma. Let ϱ be an irreducible representation of Λ in a space V, possibly infinite dimensional. Let ϱ have the minimal weight ω. Suppose that

$$\omega(\mu_\alpha) = -\tfrac{1}{2}k_\alpha \alpha(\mu_\alpha), \qquad \alpha \in \Pi,$$

where each k_α is a non-negative integer. Then if $\chi \in V$ and $\alpha \in \Pi$, there exists a finite integer N such that $\varrho(\lambda_\alpha)^N \chi = 0$.

Proof. Let v be a minimal weight vector belonging to the minimal weight ω. Let $m \geqslant 0$ and $\alpha \in \Pi$. If $\beta \in \Pi$ and $\beta \neq \alpha$, then since Π is a simple set of roots, $\alpha - \beta$ is not a root. Thus $\varrho(\lambda_{-\beta})\, \varrho(\lambda_\alpha)^m v = \varrho(\lambda_\alpha)^m \varrho(\lambda_{-\beta})\, v = 0$. If $\varrho(\lambda_{-\alpha})\, \varrho(\lambda_\alpha)^m v = 0$ also, then by Definition 8 $\varrho(\lambda_\alpha)^m v$ is a minimal weight vector. By Corollary 11, it follows that either $\varrho(\lambda_\alpha)^m\ v = 0$, or that $m = 0$. Thus

$$\text{if } \ \varrho(\lambda_\alpha)^m v \neq 0 \ \text{ and } \ m \neq 0, \ \text{ it follows that } \ \varrho(\lambda_{-\alpha})\, \varrho(\lambda_\alpha)^m v \neq 0. \tag{13}$$

We next show that if $\alpha \in \Pi$, then $\varrho(\lambda_\alpha)^N v = 0$ for sufficiently large n. To this end, suppose the contrary. Then the infinite-dimensional subspace $V^{(0)}$ of V spanned by the vectors v, $\varrho(\lambda_\alpha)\, v$, $\varrho(\lambda_\alpha)\, v, \ldots$ is invariant under λ_α and $\lambda_{-\alpha}$. If $\Lambda^{(0)}$ is the subalgebra of Λ spanned by λ_α, $\lambda_{-\alpha}$, and μ_α, and ϱ_0 is the representation of $\Lambda^{(0)}$ in $V^{(0)}$ defined by $\varrho_0(\lambda) = \varrho(\lambda)|V^{(0)}$, $\lambda \in \Lambda^{(0)}$, then ϱ_0 is seen by[2] (13) to be irreducible. ϱ_0 has the minimal weight ω. If we put $\lambda_{\pm\alpha} = \lambda_\pm$ and $\mu = \mu_\alpha/\alpha(\mu_\alpha)$, we see that $\Lambda^{(0)}$ is isomorphic to the algebra of type A_2 studied in the paragraph preceding the statement of the present lemma, and that $\omega(\mu)$ is half a non-negative integer. But therefore, by the cited paragraph, ϱ_0 must be finite dimensional, a contradiction. This proves that $\varrho(\lambda_\alpha)^N v = 0$ for N sufficiently large.

By Corollary 10 (c), every vector $x \in V$ is a linear combination of vectors $y = \varrho(\lambda_{\alpha_1}) \cdots \varrho(\lambda_{\alpha_k})\, v$, where $\alpha_1, \ldots, \alpha_k \in \Pi$. Thus, to prove the present lemma, we have only to show that there exists an N_1 so large that

$$\varrho(\lambda_\alpha)^{N_1} \varrho(\lambda_{\alpha_1}) \cdots \varrho(\lambda_{\alpha_k})\, v = 0 \tag{14}$$

Now, an evident induction using Corollary 3.8 and Lemma 3.11 will show that $\varrho(\lambda_\alpha)^{N_1} \varrho(\lambda_{\alpha_1}) \cdots \varrho(\lambda_{\alpha_k})\, v$ is a linear combination of vectors of the form

$$\varrho(\lambda_{\beta_1}) \cdots \varrho(\lambda_{\beta_i})\, \varrho(\lambda_\alpha)^p\, v,$$

where $\beta_1, \ldots, \beta_j$ are roots, $j \leqslant k$ and where $\beta_1 + \cdots + \beta_j = \alpha_1 + \cdots + \alpha_k + (N_1 - p)\alpha$. Thus $N_1 - p$ is bounded above by some integer N_2, and, if we choose $N_1 > N_2 + N$, (14) evidently follows, and our lemma is proved.

Using Lemma 21, we may now imitate the proof of Lemma 3.11, to obtain the following statement.

22. Lemma. Let ϱ be as in Lemma 21, and let $\alpha \in \Pi$. Let ω' be a weight of ϱ. Then so is

$$\omega' - \frac{2\omega'(\mu_\alpha)}{\alpha(\mu_\alpha)}\alpha.$$

Proof. Let $v' \in V$ be a weight vector belonging to the weight ω', so that by Corollary 10 (e) $\varrho(\mu)\, v' = \omega'(\mu)\, v', \mu \in M$. By the preceding lemma, the subspace V' of V spanned by all the vectors $\varrho(\lambda_{-\alpha})^p\, v$ and $\varrho(\lambda_\alpha)^p\, v'$ is finite-dimensional. The space V' is plainly invariant under the operators $\varrho(\lambda_{-\alpha})$ and $\varrho(\lambda_\alpha)$, so that $\varrho'(\lambda) = \varrho(\lambda)|V'$ defines a finite-dimensional representation of the 3-dimensional subalgebra Λ' of Λ spanned by λ_α, $\lambda_{-\alpha}$, and μ_α. Plainly, ω' is a weight of ϱ', and the weight space $V'_{\omega'+r\alpha}$ of ϱ' is spanned by a single vector $\varrho(\lambda_\alpha)^r\, v'$ or $\varrho(\lambda_{-\alpha})^{-r}\, v'$, and is therefore one-dimensional. Let R be the set of integers r for which $\omega' + r\alpha$ is a weight of ϱ'. An interval $r_- \leqslant r \leqslant r_+$ is said to be *isolated* if $r_- \in R$, $r_+ \in R$, but $r_- - 1 \notin R$, $r_+ + 1 \notin R$. If such an isolated interval is given, then, by Lemma 2.10, the space

$$V'' = \sum_{r_- \leq r \leq r_+} V'_{\omega'+r\alpha}$$

is invariant under $\varrho(\lambda_\alpha)$ and $\varrho(\lambda_{-\alpha})$. Thus

$$\operatorname{tr}(\varrho'(\mu_\alpha)|V'') = \operatorname{tr}([\varrho'(\lambda_\alpha)|V''][\varrho'(\lambda_{-\alpha})|V''] - [\varrho'(\lambda_{-\alpha}|V''][\varrho'(\lambda_\alpha)|V'']) = 0$$

But plainly

$$\operatorname{tr}(\varrho'(\mu_\alpha)|V'') = \sum_{\substack{r_- \leq r \leq r_+ \\ r \in R}} (\omega'(\mu_\alpha) + r\alpha(\mu_\alpha)).$$

This shows that the ratio $\dfrac{\omega'(\mu_\alpha)}{\alpha(\mu_\alpha)}$ is equal to the average value A_{r_-, r_+} of the integer r over the set $\{r \in R | r_- \leqslant r \leqslant r_+\}$; here, $r_- \leqslant r \leqslant r_+$ can be any isolated interval. If there were any $r^* \in R$ such that $r^* > r_+$, we could find a second isolated interval $r_- \leqslant r \leqslant r_{++}$ with $r_{++} > r_+$. But it is plain that the average value or r over this second interval would be strictly larger than A_{r_-, r_+}, which is impossible. A similar argument applies to r_-, so that we see that there can be only one isolated interval, i.e., R itself must be an interval, with end points r_- and r_+. We have then $-2\dfrac{\omega'(\mu_\alpha)}{\alpha(\mu_\alpha)} = r_+ + r_-$, and since $r_- \leqslant 0$, $-\dfrac{2\omega'(\mu_\alpha)}{\alpha(\mu_\alpha)} \leqslant r_+$, and $\omega' - \dfrac{2\omega'(\mu_\alpha)}{\alpha(\mu_\alpha)}\alpha$ is a weight. Q.E.D.

23. Corollary. Let ϱ be as in Lemma 21. Then ϱ is finite dimensional.

Proof. By the preceding lemma, if ω' is a weight of ϱ and $\alpha \in \Pi$, so is $W_\alpha\omega'$, where W_α is the *Weyl reflection* of Definition 5.8. It is plain from Definition 3.17 that $-\Pi$ is a simple set of roots. Thus, by Theorem 5.4 and Corollary 5.10, there exists a product $W = W_{\alpha_1} \cdots W_{\alpha_k}$ of Weyl reflections, $\alpha_i \in \Pi$, $i = i, \ldots, k$, such that $W\Pi = -\Pi$. If $\omega + \sum_{\alpha\in\Pi} n_\alpha\alpha$ is a weight, so is $W\omega - \sum_{\alpha\in\Pi} n_{W^{-1}\alpha}$. Then, by Corollary 10 and the linear independence of the functionals $\alpha \in \Pi$, $\sum_{\alpha\in\Pi} |n_\alpha|$ remains bounded as $\omega + \sum_{\alpha\in\Pi} n_\alpha\alpha$ varies over the set of all weights of ϱ. Using Corollary 10(f) we see that ϱ is finite dimensional. Q.E.D.

Theorem 13 is now an evident consequence of Corollaries 20 and 23.

7. The radical splitting theorem.
Existence of Lie groups with a given Lie algebra

1. Definition. Let P and Λ be given Lie algebras, and let σ be a representation of Λ on the vector space P. Suppose that $\sigma(\lambda)[\varrho_1, \varrho_2] = [\sigma(\lambda)\varrho_1, \varrho_2] + [\varrho_1, \sigma(\lambda)\varrho_2]$ for $\lambda \in \Lambda$, and $\varrho_1, \varrho_2 \in P$. The *semi-direct sum* $P \oplus_\sigma \Lambda$ of P and Λ is the Lie algebra defined as follows

(i) $P \oplus_\sigma \Lambda$, regarded as a vector space, is identical with $P \oplus \Lambda$, i.e. the set of all ordered pairs $\{\varrho, \lambda\}$ with $\varrho \in P$ and $\lambda \in \Lambda$.

(ii) The Lie product in $P \oplus_\sigma \Lambda$ is defined by

$$[\{\varrho, \lambda\}, \{\varrho', \lambda'\}] = \{[\varrho, \varrho'] + \sigma(\lambda')\varrho - \sigma(\lambda)\varrho', [\lambda, \lambda']\}.$$

The reader will easily verify that $P \oplus_\sigma \Lambda$, as defined, is in fact a Lie algebra. Our first main aim in the present section is to prove the following important structure theorem for general Lie algebras.

2. Theorem. Any Lie algebra may be written as the semi-direct sum $P \oplus_\sigma \Lambda_0$ of a solvable algebra P and a semi-simple algebra Λ_0.

To prove Theorem 2, we must first build up a certain amount of machinery.

We will call subalgebra Λ_0 of a Lie algebra Λ *solvable* if Λ_0, regarded as an algebra in its own right, is solvable (cf. Definition 1.7(b)).

3. Lemma. Let J_1 and J_2 be two solvable ideals in a Lie algebra Λ. Then $J_1 + J_2$ is a solvable ideal.

Proof. It is clear that $J_1 + J_2$ is an ideal. We have plainly $(J_1 + J_2)^{(2)} \subseteq J_1^{(2)} + J_2$, and therefore, inductively, $(J_1 + J_2)^{(n)} \subseteq J_1^{(n)} + J_2$ for all n. Since J_1 is solvable, $(J_1 + J_2)^{(n+m)} = 0$ for sufficiently large $n + m$. Q.E.D.

Since Λ is finite dimensional, the following corollary is an immediate consequence of Lemma 3.

4. Corollary. Let Λ be a Lie algebra. Then Λ contains a unique solvable ideal P of maximum dimension. This ideal contains every solvable ideal of Λ.

5. Definition. The unique maximal solvable ideal P of a Lie algebra Λ is called the *radical* of Λ.

We now repeat Definition I.6.7 for the reader's convenience.

6. Definition. Let Λ be a Lie algebra, and $J \subseteq \Lambda$ an ideal. Then the *factor algebra or quotient algebra* $M = \Lambda/J$ is the Lie algebra defined as follows:

(a) Regarded as a vector space, M is identical with the factor space Λ/J of Λ by its subspace J. Thus the elements of M are the cosets $\lambda + J$ of J.

(b) The Lie product in Λ/J is given by $[\lambda + J, \mu + J] = [\lambda, \mu] + J$.

The reader will easily verify that Λ/J satisfies all the axioms defining a Lie algebra.

7. Lemma. Let Λ be a Lie algebra, J an ideal of Λ. Then

(i) The natural map $\phi : \Lambda \to \Lambda/J$ which sends each $\lambda \in \Lambda$ into its coset $\lambda + J$ is a Lie homomorphism of Λ.

(ii) The map $J_0 \to \phi^{-1}(J_0)$ sets up a $1 - 1$ correspondence between

(a) the set of all subalgebras of Λ/J and the set of all subalgebras of Λ which contain J and

(b) the set of all ideals of Λ/J and the set of all ideals of Λ which contain J.

Proof. All these statements are elementary consequences of Definition 6; we leave it to the reader to verify them.

8. Lemma. Let Λ be a Lie algebra, and P its radical. Then Λ/P is semi-simple. Conversely, if J is an ideal of Λ, and Λ/J is semi-simple, then $J \supseteq P$.

Proof. Direct Part. If not, let J_0 be a proper solvable ideal of Λ/P. Let ϕ be as in Lemma 7. Then, by Lemma 7, $J_1 = \phi^{-1}(J)$ is a proper ideal of Λ, containing P as a proper subideal. Since ϕ is a Lie homomorphism, we have $\phi(J_1^{(n)}) = J_0^{(n)}$. Thus $\phi(J_1^{(n)}) = 0$ for sufficiently large n, i.e., $J_1^{(n)} \subseteq P$ for sufficiently large n, and thus $J_1^{(n+m)} \equiv 0$ for sufficiently large $n + m$. Thus J_1 is solvable. But this contradicts the fact that P is the large stsolvable ideal of Λ, completing the proof of the direct part of our lemma.

Converse Part. Let ϕ be the natural map of Λ onto Λ/J as in Lemma 7. Then since ϕ is a Lie homomorphism, $\phi(P)$ is an ideal. It is clear by induction that $\phi(P^{(n)}) = (\phi(P))^{(n)}$. Thus $\phi(P)$ is a solvable ideal. If Λ/J is semi-simple, it follows that $\phi(P) = \{0\}$ i.e., $J \supseteq P$. Q.E.D.

Now we begin the proof of Theorem 2. We first prove it in the following simple special case.

9. Lemma. Let Λ be a Lie algebra, P its radical. Suppose that $[\Lambda, P] = 0$. Then there exists a subalgebra Λ_0 of Λ such that $\Lambda = \Lambda_0 \oplus P$.

Proof. Let $\mathscr{T}$ be the set of linear transformations T of Λ into Λ such that $T\Lambda \subseteq P$, and such that $T|P$ is a constant multiple $\alpha(T)\, I$ of the identity. Let $\mathscr{T}_0$ be the set of all $T \in \mathscr{T}$ such that $\alpha(T) = 0$, i.e., the subspace consisting of all $T \in \mathscr{T}$ such that $T|P = 0$. Plainly, the dimension of $\mathscr{T}$ exceeds that of $\mathscr{T}_0$ by 1. Define a representation σ of Λ on $\mathscr{T}$ by putting $\sigma(\lambda)\, T = \text{Ad}\,(\lambda)\, T - T\, \text{Ad}\,(\lambda)$. Then it is easily verified that $\sigma(\lambda)\, \mathscr{T} \subseteq \mathscr{T}_0$ for $\lambda \in \Lambda$.

It follows immediately from the hypotheses of our lemma that $\sigma(\varrho) = 0$ for all $\varrho \in P$. Thus, if we define $\sigma'(\lambda + P) = \sigma(\lambda)$ for each coset $\lambda + P$ in the factor algebra $\Lambda' = \Lambda/P$, σ' is a representation of Λ' in the space $\mathscr{T}$. But, by Lemma 8, Λ/P is semi-simple. It follows by the remark following Corollary 6.6 that there exists a transformation $T^+ \in \mathscr{T}$, $T^+ \notin \mathscr{T}_0$, such that $\sigma'(\lambda + P)\, T^+ = 0$, $\lambda \in \Lambda$, i.e., such that $\text{Ad}\,(\lambda)\, T^+ = T^+\, \text{Ad}\,(\lambda)$. Thus, if we put $\Lambda_0 = \{\lambda \in \Lambda | T^+ \lambda = 0\}$, Λ_0 is easily seen to be a subalgebra of Λ, and we have $\Lambda_0 \oplus P = \Lambda$. Q.E.D.

A slight improvement of the preceding argument gives the following lemma, which is also a special case of Theorem 2.

10. Lemma. Let Λ be a Lie algebra, P its radical. Suppose that $[P, P] = 0$, but that there is no $\varrho \in P$ such that $[\varrho, \lambda] = 0$ for all Λ. Then there exists a subalgebra Λ_0 of Λ such that $\Lambda = \Lambda_0 \oplus P$.

Proof. Define $\mathscr{T}$, $\mathscr{T}_0$, and σ as in the first paragraph of the preceding proof, so that all the assertions of the said paragraph stand at our disposition. Now, every linear transformation $T = \operatorname{Ad}(\varrho)$, where $\varrho \in P$, plainly belongs to $\mathscr{T}_0$. Let $\mathscr{S}$ be the subspace of $\mathscr{T}$, consisting of transformations of this form. Since $\operatorname{Ad}(\lambda)\operatorname{Ad}(\varrho) - \operatorname{Ad}(\varrho)\operatorname{Ad}(\lambda) = \operatorname{Ad}([\lambda, \varrho])$, it is clear that $\sigma(\lambda)\mathscr{S} \subseteq \mathscr{S}$ for all $\lambda \in \Lambda$. Since

$$(\sigma(\varrho)\,T)(\lambda) = \operatorname{Ad}(\varrho)\,T\lambda - T\operatorname{Ad}(\varrho)\,\lambda = -\alpha(T)\operatorname{Ad}(\varrho)\,\lambda \quad \varrho \in P$$

so that

$$\sigma(\varrho)\,T = -\alpha(T)\operatorname{Ad}(\varrho)\,\lambda \qquad \varrho \in P \tag{1}$$

it is clear that $\sigma(\varrho)\,T \in \mathscr{S}$ if $\varrho \in P$. Let $\sigma'(\lambda)$ be the representation of Λ on the space $\mathscr{T}/\mathscr{S}$ (whose elements are the cosets $T + \mathscr{S}$ of $\mathscr{S}$) defined by $\sigma'(\lambda)(T + \mathscr{S}) = \sigma(\lambda)\,T + \mathscr{S}$. Then $\sigma'(\varrho) = 0$ for $\varrho \in P$. Thus, if we define $\sigma''(\lambda + P) = \sigma'(\lambda)$ for each coset $\lambda + P$ in the factor algebra $\Lambda' = \Lambda/P$, σ'' is a representation of Λ' in the space $\mathscr{T}/\mathscr{S}$. Plainly $\sigma''(\lambda')(\mathscr{T}/\mathscr{S}) \subseteq \mathscr{T}_0/\mathscr{S}$ for each $\lambda' \in \Lambda'$. By Lemma 8, Λ' is semi-simple. It follows by the remark following Corollary 6.6 that there exists an element $\tilde{T} \in \mathscr{T}/\mathscr{S}$, $\tilde{T} \notin \mathscr{T}_0/\mathscr{S}$, such that $\sigma''(\lambda + P)\,\tilde{T} = 0$ for each $\lambda \in \Lambda$. That is, there exists an element $T^+ \in \mathscr{T}$, $T^+ \notin \mathscr{T}_0$, such that

$$\sigma(\lambda)\,T^+ \in \mathscr{S}, \qquad \lambda \in \Lambda. \tag{2}$$

We may plainly suppose that $\alpha(T^+) = -1$. Put $\Lambda_0 = \{\lambda \in \Lambda | \sigma(\lambda)\,T^+ = 0\}$, so that since σ is a representation, Λ_0 is a subalgebra of Λ. Since by (1) $\sigma(\varrho)\,T^+ = \operatorname{Ad}(\varrho)$ for $\varrho \in P$, and since by hypothesis $\operatorname{Ad}(\varrho) \neq 0$ if $\varrho \neq 0$, it is clear that Λ_0 and P have no common point. If $\lambda \in \Lambda$ then by (2) $\sigma(\lambda)\,T^+$ has the form $\operatorname{Ad}(\varrho)$ with $\varrho \in P$; thus $\sigma(\lambda - \varrho)\,T^+ = 0$. This proves that $\Lambda = \Lambda_0 \oplus P$, and completes the proof of our lemma. Q.E.D.

Now we can easily give the proof of Theorem 2.

Proof of Theorem 2. We first prove that if Λ is a Lie algebra, and P its radical, then there exists a subalgebra Λ_0 of Λ such that $\Lambda = \Lambda_0 \oplus P$. We prove this statement by induction on the dimension n of P. If $n = 1$, our result follows either by Lemma 9 or by Lemma 10. Now inductively, suppose $n > 1$, and suppose that neither Lemma 9 nor Lemma 10 applies immediately (in the contrary case, we have nothing to prove). Then $[\Lambda, P] \neq 0$, while either $P^{(2)} \neq 0$, or there exists a non-zero $\varrho \in P$ such that $[\varrho, \lambda] = 0$ for all $\lambda \in \Lambda$. In the first case, $P^{(2)}$ is a proper ideal properly contained in P. In the second case, the set of all ϱ such that $[\varrho, \lambda] = 0$ for all $\lambda \in \Lambda$ is a proper ideal properly contained in P. Thus, in any case, there is an ideal $J \neq \{0\}$ properly contained in P.

Consider the Lie algebra $\Lambda' = \Lambda/J$, and let ϕ be the natural map of Λ onto Λ/J described by Lemma 7. Then $\phi(P)$ is an ideal of Λ'. Let $\Lambda'' = \Lambda'/\phi(P)$, and let ψ be the natural map of Λ' onto $\Lambda'/\phi(P)$, which sends each $\lambda' \in \Lambda'$ into its coset $\lambda' + \phi(P)$. Then $\psi\phi$ is a representation of Λ onto Λ'', whose kernel is precisely P. The reader will easily show from this that the Lie algebras Λ/P and Λ'' are isomorphic. Thus, by Lemma 8, $\phi(P)$ contains the radical P' of Λ'. On the other hand, it follows by induction that $\phi(P^{(n)}) = (\phi(P))^{(n)}$; thus $\phi(P)$ is solvable, and hence $\phi(P) = P'$. It is then plain that $\dim (P') < \dim (P)$. Hence, by inductive hypothesis, there exists a subalgebra Λ_0' of Λ' such that $\Lambda_0' \oplus \phi(P) = \Lambda'$. Let $\Lambda_1 = \phi^{-1}(\Lambda_0')$. Then, by Lemma 7, Λ_1 is a subalgebra of Λ containing J. We have $\Lambda_1 \cap P = J$.

Since ϕ is a representation of Λ_1 onto Λ_0' whose kernel is J, it is easily seen that Λ_1/J is isomorphic to Λ_0'. Thus, by Lemma 8, J contains the radical P_1 of Λ_1. On the other hand, $J \subseteq P$ so J is solvable. Hence $J = P_1$. Since $\dim (J) < \dim (P)$, it follows, again by inductive hypothesis, that there exists a subalgebra Λ_0 of Λ_1 such that $\Lambda_1 = \Lambda_0 \oplus J$. Since $\Lambda_1 \cap P = J$, $\Lambda_0 \cap P = \{0\}$. Plainly $\psi\phi(\Lambda_0) = \psi(\Lambda') = \Lambda/P$, so that $\Lambda_0 \oplus P = \Lambda$. This completes our proof that there exists a subalgebra Λ_0 of Λ such that $\Lambda = \Lambda_0 \oplus P$.

Now let σ denote the representation $\lambda_0 \to \mathrm{Ad}\,(\lambda_0)|P$ of Λ_0 on the vector space P. Then it is easily seen from Definition 1 that Λ is isomorphic to $P \oplus_\sigma \Lambda_0$; and Theorem 2 follows at once. Q.E.D.

11. Corollary. Every Lie algebra Λ may be built up as a repeated semi-direct sum

$$\Lambda = (((\Lambda_1 \oplus_{\sigma_1} \Lambda_2) \oplus_{\sigma_2} \Lambda_3) \oplus_{\sigma_3} \cdots (\oplus_{\sigma_n} \Lambda_{1+n}$$

in which each of the Lie algebras Λ_i is either 1-dimensional or simple.

Proof. It is clear from Theorem 2 that we have only to prove our assertion in case Λ is solvable. Now, if Λ is solvable, $\Lambda^{(2)} \subseteq \Lambda$; and then any subspace $M \subseteq \Lambda^{(2)}$ is an ideal of Λ. Choose M such that $\dim (M) + 1 = \dim (\Lambda)$; and choose $\lambda_0 \notin M$, $\lambda_0 \in \Lambda$. Then if Λ_0 denotes the 1-dimensional subalgebra of Λ spanned by λ_0, we have $\Lambda = \Lambda_0 \oplus M$. Let σ denote the representation $\lambda_0 \to \mathrm{Ad}\,(\lambda_0)|M$ of Λ_0 on the vector space M. Then it is easily seen from Definition 1 that Λ is isomorphic to $M \oplus_\sigma \Lambda_0$.

Our Corollary follows at once by induction on the dimension of Λ. Q.E.D.

We may use Corollary 11 to prove the following theorem, which completes an important piece of unfinished business.

12. Theorem. Let Λ be a real Lie algebra. Then there exists a Lie group G such that Λ is isomorphic to the Lie algebra $\Lambda(G)$ of G.

Theorem 12 is an evident consequence of Corollary 11 and of the two following Lemmas.

13. Lemma. Let G_1 and G_2 be connected Lie groups, and let Λ_1 and Λ_2 be their respective Lie algebras. Let σ be a representation of Λ_1 on the vector space Λ_2, and suppose

$$\sigma(\lambda_1)\,[\lambda_2, \lambda_2'] = [\sigma(\lambda_1)\,\lambda_2, \lambda_2'] + [\lambda_2, \sigma(\lambda_1)\,\lambda_2'], \quad \text{for} \quad \lambda_1 \in \Lambda_1,$$

$\lambda_2, \lambda_2' \in \Lambda_2$. Then the semi-direct sum $\Lambda_2 \oplus_\sigma \Lambda_1$ is isomorphic to the Lie algebra of a certain Lie group G.

14. Lemma. Let Λ be a real Lie algebra which is either 1-dimensional or simple. Then Λ is isomorphic to the Lie algebra $\Lambda(G)$ of some connected Lie group G.

Proof of Lemma 13: Since (cf. Theorem I.2.10) a Lie group and its universal covering group are locally isomorphic and hence have the same Lie algebra, we may assume without loss of generality that G_1 and G_2 are simply connected.

Since σ is a homomorphism of Λ_1 into the Lie algebra of all linear transformations of Λ_2 into itself, it follows by Theorem I.5.6 that there is a unique homomorphism σ' sending the group G_1 into the non-singular linear transformations on Λ_2. We further have $\sigma'(\exp \lambda_1) = \exp(\sigma\lambda_1)$. We now maintain that for fixed $g_1 \in G_1$ that $T = \sigma'(g_1)$ is an *automorphism* of Λ_2 into Λ_2. This may be seen by observing that for fixed $\lambda_1 \in \Lambda_1$, $\lambda_2, \mu_2 \in \Lambda_2$,

$$\sigma'(\exp t\lambda_1)\,[\lambda_2, \mu_2] \quad \text{and} \quad [\sigma'(\exp t\lambda_1)\,\lambda_2, \sigma'(\exp t\lambda_2)\,\mu_2]$$

both satisfy the same differential equation ($dv/dt = \sigma(\lambda_1)\,v$) and the same initial condition. This shows that $\sigma'(g_1)$ is an automorphism of Λ_2 if $g_1 = \exp \lambda_1$. (The inverse is clearly $\sigma'(g_1^{-1})$.) But any $g \in G_1$ may be written as a finite product of exponentials and therefore the result follows since σ' is a homomorphism of G_1.

If we fix $g_1 \in G_1$, then by the preceding paragraph $\sigma'(g_1) : \Lambda_2 \to \Lambda_2$ is an automorphism. Therefore, again by Theorem 15.6, we can find a unique

τ such that $\tau(g_1) : G_2 \to G_2$ is a homomorphism and such that $\tau(g_1) \exp \lambda_2 = \exp(\sigma'(g_1)\lambda_2)$ for all $\lambda_2 \in \Lambda_2$, $g_1 \in G_1$. Now letting g_1 vary, we have

$$\tau(g_1 h_1) \exp \lambda_2 = \exp \sigma'(g_1 h_1) \lambda_2 = \exp \sigma'(g_1)(\sigma'(h_1)\lambda_2)$$

for $\lambda_2 \in \Lambda_2$, $g_1, h_1 \in G_1$, since σ' is a homomorphism. Thus

$$\tau(g_1 h_1) \exp \lambda_2 = \tau(g_1)\tau(h_1) \exp \lambda_2 .$$

Since $\tau(g_1)$ is a homomorphism, it follows on writing $g_2 \in G_2$ as a product of exponentials that $\tau(g_1 h_1) g_2 = \tau(g_1)\tau(h_1) g_2$. Thus, $\tau(g_1 h_1) = \tau(g_1)\tau(h_1)$ and τ represents G_1 as a group of automorphisms on G_2.

We now define a *semi-direct product* G of the groups G_1 and G_2 as follows. The elements of G are the ordered pairs $\{g_2, g_1\}$ with $g_i \in G_i$. We define the operation of multiplication in G by the formula

$$\{g_2, g_1\} \cdot \{g_2', g_1'\} = \{g_2(\tau(g_1) g_2'), g_1 g_1'\}.$$

We shall verify that the Lie algebra $\Lambda(G)$ of this group is isomorphic to $\Lambda_2 \oplus_\sigma \Lambda_1$; this will plainly yield the present lemma as a consequence. We reason as follows:

First, note that it is any easy matter to verify that G is a group. In fact, the above multiplication clearly admits $\{e_2, e_1\}$ as identity (e_i the identity of G_i) and admits an inverse given by

$$\{g_2, g_1\}^{-1} = \{\tau(g_1^{-1}) g_2^{-1}, g_1^{-1}\}.$$

The associative law is easily verified. Moreover, it is seen by the above formulas that if G is given the C^∞ structure of the product manifold of G_1 and G_2, then G is a Lie group. For it is clear that $\tau(g_1) g_2'$ is C^∞ as a mapping of G into G_2.

The mappings $g_1 \to \{e_2, g_1\}$ and $g_2 \to \{g_2, e_1\}$ are regular embeddings of G_i in G. We may thus write $G_i \subseteq G$ and consider G_i as a subgroup of G. Since the embedding is regular, we have $\Lambda_i \subseteq \Lambda$ in a natural way, where Λ is the Lie algebra of G. From the general theory of manifolds, we have $\Lambda = \Lambda_2 \oplus \Lambda_1$ (as a vector space), where we have identified $\lambda_1 \in \Lambda_1$ with $\{0, \lambda_1\}$ and $\lambda_2 \in \Lambda_2$ with $\{\lambda_2, 0\}$. Since the embedding of G_i in G is a homomorphism, we see that Λ_i is embedded in Λ as a *subalgebra*. It is therefore only necessary to compute $[\lambda_1, \lambda_2]$ (where $\lambda_i \in \Lambda_i$) in order to determine the algebraic structure of Λ completely.

We first rewrite the definition of multiplication with the help of our identification $G_i \subseteq G$. This yields

$$g_1 g_2 = \tau(g_1) g_2 \cdot g_1 \quad \text{or} \quad g_1 g_2 g_1^{-1} = \tau(g_1) g_2 .$$

Now write $g_2 = \exp t\lambda_2$ and take derivatives at $t = 0$. This gives (see Definition 13.1):

$$g_1\lambda_2 g_1^{-1} = \frac{d}{dt}\,\tau(g_1)\exp t\lambda_2\bigg|_{t=0}$$

$$= \frac{d}{dt}\exp\,(t\sigma'(g_1)\,\lambda_2)\bigg|_{t=0}$$

$$g_1\lambda_2 g_1^{-1} = \sigma'(g_1)\,\lambda_2\,.$$

Now put $g_1 = \exp t\lambda_1$ and take derivatives again. Using Definition 13.7, we have

$$[\lambda_1, \lambda_2] = \frac{d}{dt}\,\sigma'\,(\exp t\lambda_1)\,\lambda_2\bigg|_{t=0}$$

$$= \frac{d}{dt}\exp\,(t\sigma(\lambda_1))\,\lambda_2\bigg|_{t=0}$$

$$[\lambda_1, \lambda_2] = \sigma(\lambda_1)\,\lambda_2\,.$$

From this, we obtain

$$\begin{aligned}[\{\lambda_2, \lambda_1\}, \{\lambda_2', \lambda_1'\}] &= [\lambda_2 + \lambda_1, \lambda_2' + \lambda_1'] \\ &= [\lambda_2, \lambda_2'] + [\lambda_1, \lambda_2'] - [\lambda_1', \lambda_2] + [\lambda_1, \lambda_1'] \\ &= [\lambda_2, \lambda_2'] + \sigma(\lambda_1)\,\lambda_2' - \sigma(\lambda_1')\,\lambda_2 + [\lambda_1, \lambda_1'] \\ &= \{[\lambda_2, \lambda_2'] + \sigma(\lambda_1)\,\lambda_2' - \sigma(\lambda_1')\,\lambda_2, [\lambda_1, \lambda_1']\},\end{aligned}$$

and therefore, $\Lambda = \Lambda_2 \oplus_\sigma \Lambda_1$ which is the result. Q.E.D.

Proof of Lemma 14. The 1-dimensional Lie algebra is the Lie algebra of the additive group of real numbers. Next, let Λ be simple. Then since the set of $\lambda \in \Lambda$ such that $[\lambda, \mu] = 0$ for all $\mu \in \Lambda$ forms an ideal this set contains no non-zero λ. Hence the map $\lambda \to \mathrm{Ad}\,(\lambda)$ is an isomorphism of Λ onto a subalgebra Λ' of the Lie algebra Λ^* of all linear transformations in Λ. The algebra Λ^* is, by Lemma I.3.15, the Lie algebra of the group G^* of all non-singular linear transformations in Λ. By Theorem I.6.3, there exists a Lie subgroup G' of G whose Lie algebra is isomorphic to Λ'. This proves Lemma 14. As we have already noted, Theorem 12 follows. Q.E.D.

Note that our proof of Theorem 12 gives a very explicit recipe for the construction of the group G whose Lie algebra is Λ.

We may also use Theorem 2 to describe all the irreducible representations of a general complex Lie algebra. Unfortunately, since Theorem 6.2 cannot be generalized to arbitrary Lie algebras, such a result is not particularly decisive. The theorem in question is as follows.

15. Theorem. Let P and Λ_0 be complex Lie algebras, and let σ be a representation of Λ_0 on the vector space P. Suppose that P is solvable, and that Λ_0 is semi-simple. Suppose also that $\sigma(\lambda)\,[\varrho_1, \varrho_2] = [\sigma(\lambda)\,\varrho_1, \varrho_2] + [\varrho_1, \sigma(\lambda)\,\varrho_2]$ for $\lambda \in \Lambda$ and $\varrho_1, \varrho_2 \in P$. Let $\Lambda = P \oplus_\sigma \Lambda_0$. Let τ be an irreducible representation of Λ on a finite dimensional vector space V. Then there exists an irreducible representation τ_0 of Λ_0 on V, and a linear functional α defined on P and satisfying $\alpha([\varrho_1, \varrho_2]) = 0$ and $\alpha(\sigma(\lambda)\,\varrho_1) = 0$ for $\lambda \in \Lambda$ and $\varrho_1, \varrho_2 \in P$, such that

$$\tau\{\varrho, \lambda\} = (\alpha(\varrho)\, I + \tau_0(\lambda)) \qquad \varrho \in P, \lambda \in \Lambda_0. \tag{3}$$

Conversely, if α and τ_0 are given and have the above properties, (1) defines an irreducible representation of $P \oplus_\sigma \Lambda_0$.

Proof. First suppose that α and τ_0 are given, and that τ is defined by (1). We have

$$[\{\varrho, \lambda\}, \{\varrho', \lambda'\}] = \{[\varrho, \varrho'] + \sigma(\lambda')\,\varrho - \sigma(\lambda)\,\varrho', [\lambda, \lambda']\} \quad \text{by Definition 1.}$$

Thus

$$\tau([\{\varrho, \lambda\}, \{\varrho', \lambda'\}]) = \tau_0([\lambda_1 \lambda']).$$

Moreover

$$[\tau\{\varrho, \lambda\}, \tau\{\varrho', \lambda'\}] = [\alpha(\varrho)\, I + \tau_0(\lambda), \alpha(\varrho')\, I + \tau_0(\lambda')] = \tau_0([\lambda, \lambda']).$$

This shows that τ is a representation of Λ. Since τ_0 is irreducible, so is τ.

Conversely, let τ be an irreducible representation of $\Lambda = P \oplus_\sigma \Lambda_0$. Identifying $\{\varrho, 0\}$ with ϱ and $\{0, \lambda_0\}$ with λ_0, we may assume P and Λ_0 to be subalgebras of Λ. Since P is solvable, it follows by Theorem 1.9 that there exists a vector $v \neq 0$ in V and a linear functional α on P such that $\tau(\varrho)\, v = \alpha(\varrho)\, v$, $\varrho \in P$. It is plain that $P \subseteq \Lambda$ is an ideal, and therefore $[\tau(\lambda), \tau(\varrho)] \in \tau(P)$ if $\lambda \in \Lambda$ and $\varrho \in P$. Thus, for each real t we have $\exp\,(t \operatorname{Ad}\,(\tau(\lambda)))\,\tau(\varrho) \in \tau(P)$ if $\lambda \in \Lambda$ and $\varrho \in P$. Using Lemma I.5.3, it follows that for each $-\infty < t < \infty$ there exists a $\varrho'(t, \lambda, \varrho) \in P$ such that

$$\exp(-t\tau(\lambda))\,\tau(\varrho)\exp(t\tau(\lambda)) = \tau(\varrho'(t, \lambda, \varrho)).$$

Hence

$$\exp(-t\tau(\lambda))\,\tau(\varrho)\exp(t\tau(\lambda))\,v = \alpha(\varrho'(t, \lambda, \varrho))\,v,$$

i.e.

$$\tau(\varrho)\exp(t\tau(\lambda))\,v = \alpha(\varrho'(t, \lambda, \varrho))\exp(t\tau(\lambda))\,v. \tag{4}$$

This shows that for $\lambda \in \Lambda$ and $\varrho \in P$, the non-zero vector $\exp(t\varrho(\lambda))\,v$ is always an eigenvector of the linear transformation $\tau(\varrho)$. Now, the various eigenspaces of $\tau(\varrho)$ are closed sets having no vector but zero in common; hence a continuous curve cannot pass from one eigenspace to another without passing through zero. It follows that $\exp(t\tau(\lambda))\,v$ belongs to the eigenvalue $\alpha(\varrho)$ of $\tau(\varrho)$ for all real t. Differentiating with respect to t and putting $t = 0$, we find that the space of vectors v satisfying $\tau(\varrho)\,v = \alpha(\varrho)\,v$ is invariant under each operator $\tau(\lambda)$, $\lambda \in \Lambda$. Since V is irreducible for τ, we must have $\tau(\varrho)\,v = \alpha(\varrho)\,v$ for every $v \in V$. Thus, if τ_0 is the restriction of τ to the subalgebra Λ_0 of Λ, we have

$$\tau(\{\varrho, \lambda\}) = \tau(\{\varrho, 0\} + \{0, \lambda\}) = (\alpha(\varrho)\,I + \tau_0(\lambda)).$$

It is then plain that if τ is to be irreducible, so must τ_0 be irreducible. Q.E.D.

Though to do so is somewhat apart from our main purposes, let us digress to comment somewhat more extensively on solvable Lie algebras. The fact that any solvable Lie algebra Λ may be written as an iterated semi-direct sum

$$\Lambda = (\cdots((\Lambda_1 \oplus_{\sigma_1} \Lambda_2) \oplus_{\sigma_2} \Lambda_3) \oplus \cdots) \oplus_{\sigma_{n-1}} \Lambda_n,$$

where all the Lie algebras Λ_j are one-dimensional, might suggest that the classification of these algebras ought to be easy. Unfortunately, such optimism is unjustified. Take, for example, the five-dimensional Lie algebra N with basis elements $\lambda_\pm$ and μ_1, μ_2, μ_3 satisfying $[\lambda_\pm, \mu_i] = \pm\mu_i$, $i = 1, 2$; $[\lambda_\pm, \mu_3] = \mu_3$; $[\lambda_+, \lambda_-] = 0$, and $[\mu_i, \mu_j] = 0$, $i, j = 1, 2, 3$. If V is an arbitrary finite-dimensional vector space, and $T_\pm$ a pair of commuting nilpotent linear operators in V, then

$$\tau : \lambda_\pm, \mu_1, \mu_2, \mu_3 \to T_\pm, 0, 0, 0$$

defines a representation of N in V. If we regard V as a Lie algebra in which all Lie products are zero, we may introduce the semi-direct sum $V \oplus_\tau \Lambda$, which is easily seen to be a nilpotent Lie algebra. Suppose that $S_\pm$ is a second pair of commuting nilpotent linear operators in V so that

$$\sigma : \lambda_\pm, \mu_1, \mu_3, \mu_3 \rightarrow S_\pm, 0, 0, 0$$

defines a second representation of Λ in V, and consider the semidirect sum $V \oplus_\sigma \Lambda$. When are the two nilpotent Lie algebras $V \oplus_\sigma \Lambda$ and $V \oplus_\tau \Lambda$ isomorphic? This question is readily answered as follows. In the first place, it is easily verified that any element $\alpha = \{v, \nu\}$ of the semi-direct sum $V \oplus_\tau \Lambda$ such that $\mathrm{Ad}(\alpha)$ has $+1$ as an eigenvalue to which there belongs a three-dimensional eigenspace must have the form $\alpha = \{v, z_1\mu_1 + z_2\mu_2 + z_3\mu_3 + \lambda_+\}$. Similarly, if $\beta = \{v, \nu\}$ is an element of $V \oplus_\tau \Lambda$ such that $\mathrm{Ad}(\beta)$ has -1 and $+1$ as eigenvalues, $+1$ corresponding to a one-dimensional eigenspace and -1 corresponding to a 2-dimensional eigenspace, then β must have the form $\beta = \{\hat{v}, \hat{z}_1\mu_1 + \hat{z}_2\mu_2 + \hat{z}_3\mu_3 + \lambda_-\}$. The eigenspace M of α belonging to the eigenvalue $+1$ is the linear span of μ_1, μ_2 and μ_3; the intersection of the set of all elements $\gamma \in V \oplus_\tau \Lambda$ such that $\mathrm{Ad}(\alpha)^K \gamma = \mathrm{Ad}(\beta)^K \gamma = 0$ for sufficiently large K with the set $\{\gamma \in V \oplus_\tau \Lambda \mid [\gamma, \mu] = 0 \text{ for all } \mu \in M\}$ is the space $V \subseteq V \oplus_\tau \Lambda$ of γ of the form $\gamma = \{v, 0\}$. If A is any isomorphism mapping $V \oplus_\tau \Lambda$ onto $V \oplus_\sigma \Lambda$, it follows that A must map V into itself, and must have the additional property that

$$A\{0, \lambda_\pm\} = \{v_\pm, \mu_\pm + \lambda_\pm\},$$

where $\mu_\pm \in M$. But then, if $v \in V$, we have

$$AT_\pm v = A[\lambda_\pm, v] = [A\lambda_\pm, Av] = [\lambda_\pm, Av] = S_\pm Av,$$

so that the restriction A_0 of A to V satisfies $A_0 T_+ A_0^{-1} = S_+$. Conversely, if there exists a non-singular linear transformation A_0 such that $A_0 T_\pm A_0^{-1} = S_\pm$, then it is plain that $V \oplus_\tau \Lambda$ and $V \oplus_\tau \Lambda$ are isomorphic.

To classify solvable, or even nilpotent, Lie algebras up to isomorphism one would therefore have to classify pairs of commuting nilpotent matrices (T_+, T_-) up to equivalence!

$$(T_+, T_-) \sim (S_+, S_-) \tag{5}$$

if there exists a non-singular A_0 such that

$$S_\pm = A_0 T_\pm A_0^{-1}.$$

A similar construction, which we leave to the reader to elaborate, shows that to classify nilpotent Lie algebras up to isomorphism one would have to classify triples of commuting nilpotent matrices (T_1, T_2, T_3) up to equivalence:

$$(T_1, T_2, T_3) \sim (S_1, S_2, S_3) \tag{6}$$

if there exists a non-singular A_0 such that

$$S_i = A_0 T_i A_0^{-1}, \qquad i = 1, 2, 3.$$

Now these problems, which must of course be faced in any attempt to found a systematic "radical algebra", are quite complicated. This is most easily seen in the case of the second problem: if we let the matrices of the transformations T_i and S_i be $2n \times 2n$ matrices partitioned into $n \times n$ matrices as

$$T_i = \begin{pmatrix} 0 & 0 \\ t_i & 0 \end{pmatrix}, \quad S_i = \begin{pmatrix} 0 & 0 \\ s_i & 0 \end{pmatrix}, \quad i = 1, 2, 3,$$

and take $t_1 = s_1 = I$, then any A_0 such that $A_0 T_1 A_0^{-1} = S_1$ must have the form

$$A_0 = \begin{pmatrix} a_0 & 0 \\ b & a_0 \end{pmatrix}.$$

But then if $A_0 T_i A_0^{-1} = S_i$, $i = 2, 3$, we must have $a_0 t_i a_0^{-1} = s_i$, $i = 2, 3$. Thus to solve the second of the two problems above one would have to classify *arbitrary* pairs (t_2, t_3) of matrices up to similarity in the sense of (5).

We may make yet another remark. If T_+ and T_- above happen to be non-zero matrices standing in the relation $T_+ = cT_-$, then the constant c is evidently an invariant of the equivalence relation (5); thus, in contradistinction to semi-simple Lie algebras, nilpotent Lie algebras may depend upon one or several continuous parameters. In particular, while there are only countably many non-isomorphic semi-simple algebras, the cardinality of the set of non-isomorphic nilpotent Lie algebras is that of the continuum.

Given a complex Lie algebra Λ_1, we can define its *complex conjugate Lie algebra* to be the Lie algebra whose elements are the same as the elements of Λ_1, and which are added and combined together in Lie brackets in the same way as the elements of Λ_1, but with the product of an element λ_1 of Λ_1 and a complex number z redefined as $\bar{z}\lambda_1$ rather than $z\lambda_1$. It is then plain that if $S_\pm = \bar{T}_\pm$ (the bar denoting the complex conjugate matrix), then the complex conjugate of the Lie algebra $V \oplus_\tau \Lambda$ is isomorphic to the Lie algebra $V \oplus_\sigma \Lambda$. If we now let T_+ and T_- stand in the relation $T_+ = cT_-$ with c non-real, it is clear from what we have said above that $V \oplus_\tau \Lambda$ is not isomorphic to its complex conjugate. There is in particular no basis λ_i, $i = 1, \ldots, n$ of this Lie algebra for which $[\lambda_i, \lambda_j] = \sum c_{ij}^k \lambda_k$ with real constants c_{ij}^k, which shows that the property of semi-simple Lie algebras expressed by Theorem 4.5 is not shared by nilpotent Lie algebras.

PART III

Real Semi—Simple Lie Algebras

1. Structure and Representations of Simple Real Algebras

In sections II.1–II.4 of the present work, we gave a complete account of the structure of all *complex* semi-simple Lie algebras. In the present section, we intend to give a correspondingly complete analysis of the structure of real simple (and hence of the structure of real semi-simple) Lie algebras Λ. Our method will be to study the relation between Λ and its complexification $\hat{\Lambda}$, applying the preceding theory to $\hat{\Lambda}$. Our first lemma shows that the complexification $\hat{\Lambda}$ of a real simple Lie algebra Λ can never be very far from simple.

1. Lemma. Let Λ be a real simple Lie algebra, and $\hat{\Lambda}$ its complexification (cf. Definition II.1.2). The Lie algebra $\hat{\Lambda}$ is either simple or the sum of two isomorphic complex simple Lie algebras.

Proof. The algebra $\hat{\Lambda}$ consists of all formal sums $\lambda + i\mu$, $\lambda, \mu \in \Lambda$, and thus admits the conjugation $C : \lambda + i\mu \to \lambda - i\mu$. This conjugation satisfies $C^2 = I$, $C[\hat{\lambda}, \hat{\mu}] = [C\hat{\lambda}, C\hat{\mu}]$ if $\hat{\lambda}, \hat{\mu} \in \Lambda$, and $C(\hat{\lambda} + \hat{\mu}) = C\hat{\lambda} + C\hat{\mu}$, but we have $Cz\hat{\lambda} = zC\hat{\lambda}$ if z is a complex number.

By Corollary II.2.14 and Theorem II.2.17, we may write $\hat{\Lambda} = J_1 \oplus \cdots \oplus J_n$, where each J_j is an ideal of $\hat{\Lambda}$ which, regarded as a Lie algebra in its own right, is simple. Plainly then the ideals CJ_j have the same property. Hence, if $CJ_j \cap J_k \neq \{0\}$, $CJ_j = J_k$. But $[CJ_j, J_k] \subseteq CJ_j \cap J_k$, and since $[CJ_j, \hat{\Lambda}] \neq 0$ and $\hat{\Lambda} = J_1 \oplus \cdots \oplus J_n$, we must have $CJ_j \cap J_k \neq \{0\}$ for at least one k. It follows that the ideals J_j can be enumerated in such a way that $CJ_{2j-1} = J_{2j}$ and $CJ_{2j} = J_{2j-1}$ for $j = 1, \ldots, m \leqslant n/2$, and $CJ_j = J_j$ for $2m < j \leqslant n$.

It is clear that Λ is isomorphic to the real Lie algebra

$$\hat{\Lambda}_C = \{\hat{\lambda} \in \hat{\Lambda} | C\hat{\lambda} = \hat{\lambda}\}.$$

But evidently

$$\hat{\Lambda}_C = \sum_{j=1}^{m} \oplus \{\hat{\lambda} \in J_{2j-1} \oplus J_{2j} | C\hat{\lambda} = \hat{\lambda}\} \oplus \sum_{j=2m+1}^{n} \{\hat{\lambda} \in J_j | C\hat{\lambda} = \hat{\lambda}\}, \quad (1)$$

and each of the terms in this direct sum is a non-zero ideal of $\hat{\Lambda}_C$. Thus there can be at most one term in the direct sum (1), and either $n = 1$ or $n = 2$.

If $n = 1$, we have nothing to prove. If $n = 2$, we must prove that the Lie algebras J_1 and J_2 are isomorphic. This we do as follows. The Lie algebra J_1 is simple. Thus, by Theorem II.4.5, we can find a basis $\{\lambda_i\}$ for J_1 such that $[\lambda_i, \lambda_j] = \sum_k c_{i\,j}^{k} \lambda_k$, where all the coefficients $c_{i\,j}^{k}$ are rational. In terms of this basis, define a transformation C_1 of J_1 into itself by $C_1 (\sum z_i \lambda_i) = \sum \bar{z}_i \lambda_i$. Then plainly CC_1 is an isomorphism of J_1 onto CJ_1, and the proof of the present lemma is complete. Q.E.D.

2. Definition. A *conjugation* in a complex Lie algebra Λ_1 is a mapping C of Λ_1 into itself such that $C[\lambda_1, \mu_1] = [C\lambda_1, C\mu_1]$, $C^2 = I$, $C(\lambda_1 + \mu_1) = C\lambda_1 + C\mu_1$, and $C(z\lambda_1) = \bar{z}C\lambda_1$ for $\lambda_1, \mu_1 \in \Lambda_1$, and z a complex number. In the final paragraph of the proof of Lemma 1 we note the following fact, appropriately formulated in terms of the preceding definition.

3. Corollary. A complex simple Lie algebra always admits a conjugation.

If we sharpen the argument used to prove Lemma 1 ever so slightly, we can settle one of the two cases admitted by Lemma 1 completely.

4. Lemma. A simple real algebra Λ is isomorphic to either

(a) The simple real algebra Λ_2 obtained from a simple complex algebra Λ_1 by restricting the field of scalars from the complex field to the real field, or

(b) The simple real algebra $\hat{\Lambda}_C = \{\hat{\lambda} \in \hat{\Lambda} | C\hat{\lambda} = \hat{\lambda}\}$, where C is a conjugation in the simple complex algebra $\hat{\Lambda}$.

In case (a), the complexification of Λ_2 is isomorphic to $\Lambda_1 \oplus \Lambda_1$; in case (b), the complexification of $\hat{\Lambda}_c$ is isomorphic to $\hat{\Lambda}$.

Moreover, each of these two constructions does yield a simple real algebra.

Proof. If we examine the proof of Lemma 1, we see that to prove our first assertion we have only to verify that if $\hat{\Lambda} = J \oplus CJ$, where J is an ideal of $\hat{\Lambda}$ and a simple algebra in its own right, then the real Lie algebra $\hat{\Lambda}_c = \{\hat{\lambda} \in \hat{\Lambda} | C\hat{\lambda} = \hat{\lambda}\}$ is isomorphic to the algebra obtained from J by restricting the field of scalars from the complex field to the real field. But, since $C^2 = I$, $\hat{\Lambda}_c$ evidently consists of all the elements $\lambda + C\lambda$, where $\lambda \in J$, and the correspondence $\lambda \leftrightarrow \lambda + C\lambda$ between $\hat{\Lambda}_c$ and J gives the desired isomorphism.

To prove our second assertion, we argue as follows. Suppose that Λ_1 is a simple complex Lie algebra, but that the real Lie algebra Λ_2 derived from Λ_1 by restricting the field of scalars from the complex field to the real field is not simple. Using Definition II.1.2, we see immediately that the complexification $\hat{\Lambda}_2$ of Λ_2 is the direct sum of two isomorphic copies of Λ_1, each of which is an ideal in $\hat{\Lambda}_2$. Thus, by Theorem II.2.17, $\hat{\Lambda}_2$ is semi-simple, so, by Corollary II.2.14, Λ_2 is semi-simple. Using Theorem II.2.17 once more, we see that we may write Λ_2 as a direct sum $\Lambda_2 = \Lambda_2^{\{1\}} \oplus \cdots \oplus \Lambda_2^{\{n\}}$ of ideals, each of which is a real simple algebra in its own right. If $\lambda \in \Lambda_2$, then $i\lambda \in \Lambda_2$ also; from which it is plain that each of the ideals $\Lambda_2^{\{j\}}$ contains $i[\Lambda_2^{\{j\}}, \Lambda_2^{\{j\}}]$. But, since $\Lambda_2^{\{j\}}$ is simple, $\Lambda_2^{\{j\}} = [\Lambda_2^{\{j\}}, \Lambda_2^{\{j\}}]$, and thus $i\Lambda_2^{\{j\}} \subseteq \Lambda_2^{\{j\}}$, which shows that each of the sets $\Lambda_2^{\{j\}}$ is actually an ideal of the *complex* Lie algebra Λ_1. Since Λ_1 is simple, this is impossible, and from this contradiction we conclude that Λ_2 is simple.

Next suppose that $\hat{\Lambda}$ is a simple complex algebra and C a conjugation in $\hat{\Lambda}$, but that $\hat{\Lambda}_c = \{\lambda \in \hat{\Lambda} | \lambda = C(\lambda)\}$ is not simple. If $\lambda, \mu \in \hat{\Lambda}_c$ so that $\lambda + i\mu$ may be regarded either as an element of the complexification Λ_1 of $\hat{\Lambda}_c$ or as an element of $\hat{\Lambda}$, define $\tau(\lambda + i\mu) = \lambda + i\mu \in \hat{\Lambda}$, so that τ maps Λ_1 into $\hat{\Lambda}$. Plainly, τ is an isomorphism onto. Since each $v \in \hat{\Lambda}$ may be written as $v = \frac{1}{2}(v + Cv) + \frac{1}{2}i\,(v - Cv)$, τ is an isomorphism onto. By Corollary II.2.14 then, $\hat{\Lambda}_c$ is semi-simple, so that by Theorem II.2.17 we may write $\hat{\Lambda}_c$ as a direct sum $\hat{\Lambda}_c = \Lambda^{\{1\}} \oplus \cdots \oplus \Lambda^{\{n\}}$ of ideals, each of which, regarded as a Lie algebra in its own right, is simple. It is then plain from Definition II.1.2 that the complexification Λ_1 of $\hat{\Lambda}_c$ is the direct sum of the complexification of the various $\Lambda^{\{j\}}$, each of which is an ideal in Λ_1. Since Λ_1 has been shown to be simple, $n > 1$ is impossible, and thus $\Lambda_c = \Lambda^{\{1\}}$ is simple. Q.E.D.

5. Corollary. (a) Let Λ_1' and Λ_2' be two simple complex algebras, and let Λ_1 and Λ_2 be the simple real algebras obtained from these by restricting the field of scalars from the complex field to the real field. Then if Λ_1 and Λ_2 are isomorphic, Λ_1' and Λ_2' are isomorphic.

(b) No simple real algebra arising under case (a) of Lemma 4 is isomorphic to a simple real algebra arising under case (b) of Lemma 4.

Proof. We saw in the course of the preceding proof that the complexifications of the simple real algebras arising under case (b) of the preceding lemma were themselves simple, while the complexifications of the simple real algebras arising under case (a) were the direct sum of two ideals. Q.E.D.

Lemma 4 and Corollary 5 show that the simple real algebras Λ whose complexifications $\hat{\Lambda}$ are not simple are in natural one-to-one correspondence with the classes of simple complex algebras described by Theorem II.3.36; and nothing more need be said about this class of simple real algebras.

We shall therefore be interested, for the remainder of the present section, in real Lie algebras Λ defined by the formula $\Lambda = \{\hat{\lambda} \in \hat{\Lambda} | C\hat{\lambda} = \hat{\lambda}\}$, where $\hat{\Lambda}$ is a complex simple Lie algebra, and where C is a conjugation in $\hat{\Lambda}$.

The following lemma shows how close the connection between $\hat{\Lambda}$ and C really is.

6. Lemma. Let C_1 and C_2 be two conjugations in $\hat{\Lambda}$, and let

$$\Lambda_i = \{\hat{\lambda} \in \hat{\Lambda} | C_i\hat{\lambda} = \hat{\lambda}\}, \qquad i = 1, 2.$$

Then the real Lie algebras Λ_1 and Λ_2 are isomorphic if and only if there exists an automorphism τ of $\hat{\Lambda}$ such that $C_2 = \tau C_1 \tau^{-1}$.

Proof. First suppose that $C_2 = \tau C_1 \tau^{-1}$. Then plainly $\hat{\lambda} \in \Lambda_2$ if and only if $\tau^{-1}\hat{\lambda} \in \Lambda_1$; thus $\tau | \Lambda_1$ is an isomorphism of Λ_1 onto Λ_2.

Suppose conversely that there exists an isomorphism T of Λ_1 onto Λ_2. Put

$$\tau\hat{\lambda} = T\left(\frac{1}{2}(\hat{\lambda} + C_1\hat{\lambda})\right) + iT\left(\frac{1}{2i}(\hat{\lambda} - C_1\hat{\lambda})\right).$$

We easily verify that $\tau(\hat{\lambda}_1 + \hat{\lambda}_2) = \tau\hat{\lambda}_1 + \tau\hat{\lambda}_2$; $\tau(x\hat{\lambda}_1) + x\tau(\hat{\lambda}_1)$ if x is a real number; and $\tau(i\hat{\lambda}) = i\tau\hat{\lambda}$; thus τ is a linear transformation of $\hat{\Lambda}$ onto itself. We have $\tau[\hat{\lambda}_1, \hat{\lambda}_2] = [\tau\hat{\lambda}_1, \tau\hat{\lambda}_2]$ if $\hat{\lambda}_1, \hat{\lambda}_2 \in \Lambda_1$; since Λ_1 spans $\hat{\Lambda}$, this equation holds identically, and τ is an automorphism of $\hat{\Lambda}$. If $\tau C_1 \tau^{-1}\hat{\lambda} = \hat{\lambda}$, then $\tau^{-1}\hat{\lambda} \in \Lambda_1$, so that $T\tau^{-1}\hat{\lambda} = \tau\tau^{-1}\hat{\lambda} = \hat{\lambda} \in \Lambda_2$, and therefore $\tau C_1 \tau^{-1}\hat{\lambda} = \hat{\lambda} = C_2\hat{\lambda}$. Thus $C_2^{-1}\tau C_1 \tau^{-1}\hat{\lambda} = \hat{\lambda}$ if $\hat{\lambda} \in \tau\Lambda_1 = \Lambda_2$. Since $C_2^{-1}\tau C_1 \tau^{-1}$ is a linear transformation in $\hat{\Lambda}$, and since Λ_2 spans $\hat{\Lambda}$, we have therefore $C_2^{-1}\tau C_1 \tau^{-1} = I$, and hence $\tau C_1 \tau^{-1} = C_2$. Q.E.D.

7. Definition. Two conjugations C_1 and C_2 in $\hat{\Lambda}$ are *equivalent* if there exists an automorphism τ of $\hat{\Lambda}$ such that $C_2 = \tau C_1 \tau^{-1}$. The family of all conjugations in $\hat{\Lambda}$ which are equivalent to C_1 will occasionally be called the *equivalence class* of C_1. The real subalgebra Λ of $\hat{\Lambda}$ *defined by the conjugation* C_1 is $\Lambda = \{\hat{\lambda} \in \hat{\Lambda} | C\hat{\lambda} = \hat{\lambda}\}$.

In this terminology we may restate Lemma 6 as: *Two conjugations define isomorphic real subalgebras if and only of they are equivalent.*

While the preceding results set the scene, they are merely preparatory. We now begin in earnest to study the conjugations in a simple complex Lie algebra $\hat{\Lambda}$. Our first aim is to prove the existence of a conjugation with an important special property.

8. Definition. Let $\hat{\Lambda}$ be a complex Lie algebra, (λ, μ) the natural bilinear form in $\hat{\Lambda}$, and C a conjugation in $\hat{\Lambda}$. Then

(a) C is called *positive* if $-(\lambda, C\lambda) > 0$ for every non-zero $\lambda \in \hat{\Lambda}$. If C is a positive conjugation, we write $\langle \lambda, \mu \rangle_c = -(\lambda, C\mu)$, so that $\langle \lambda, \mu \rangle_c$ is a positive definite hermitian form on $\hat{\Lambda}$.

(b) If ϕ is an automorphism of $\hat{\Lambda}$, we denote the automorphism $C\phi C$ by the symbol ϕ_c.

9. Lemma. Let $\hat{\Lambda}$ be a simple complex algebra. Then there exists a positive conjugation C in $\hat{\Lambda}$.

To the proof if Lemma 9 we prefix the following preparatory lemma:

10. Lemma. Let Λ_1 be a semi-simple Lie algebra, and let Λ_2 be a simple subalgebra of Λ_1. Let $(\lambda, \mu)_1$ and $(\lambda, \mu)_2$ be the natural bilinear forms of Λ_1 and Λ_2 respectively. Then the restriction of $(\lambda, \mu)_1$ to Λ_2 is a constant multiple $a(\lambda, \mu)_2$ of $(\lambda, \mu)_2$.

Moreover, if Λ_1 and Λ_2 are complex algebras, if a Cartan subalgebra M_1 of Λ_1 contains a Cartan subalgebra M_2 of Λ_2, and if there exists a non-zero vector $\mu \in M_2$ such that every root α of M_1 in Λ_1 assumes a real value on μ, then the constant a must be positive.

Proof. We have (cf. Corollary II.2.16) $(\mathrm{Ad}(\lambda)\mu, \nu)_1 = -(\mu, \mathrm{Ad}(\lambda)\nu)_1$ and $(\mathrm{Ad}(\lambda)\mu, \nu)_2 = -(\mu, \mathrm{Ad}(\lambda)\nu)_2$ for $\lambda, \mu, \nu \in \Lambda_2$. The form $(\lambda, \mu)_2$ is non-singular by Theorem 2.3, and thus for each $\nu \in \Lambda_2$ the linear form $(\mu, \nu)_1$ on Λ_2 may be represented as $(\mu, \nu')_2$, where $\nu' \in \Lambda_2$. In virtue of the non-singularity of $(\lambda, \mu)_2$, it follows that $\nu \to \nu'$ defines a linear transformation T of Λ_2 onto itself; we have then

$$(T\lambda, \mu)_2 = (\lambda, \mu)_1, \; \lambda, \mu \in \Lambda_2.$$

Therefore

$$(\mathrm{Ad}\,(\nu)\,(T\lambda), \mu)_2 = -(T\lambda, \mathrm{Ad}\,(\nu)\,\mu)_2 = -(\lambda, \mathrm{Ad}\,(\nu)\,\mu)_1 = (\mathrm{Ad}\,(\nu)\,\lambda, \mu)_1$$
$$= (T\,\mathrm{Ad}\,(\nu)\,\lambda, \mu)_2 \quad \text{for} \quad \lambda, \mu, \nu \in \Lambda_2 .$$

Thus T commutes with all the operators Ad (ν), $\nu \in \Lambda_2$. Since Λ_2 is simple, the family of operators Ad (ν) admits no proper invariant subspace. Since every eigenspace $\{\lambda | T\lambda = z\lambda\}$ of T is such an invariant subspace, T must be a constant multiple of the identity. This proves our first assertion.

To prove the second assertion, we note that since Λ_1 (resp. Λ_2) is the direct sum of the various root spaces of M_1 in Λ_1 (resp. M_2 in Λ_2), the eigenvalues of the operator Ad (μ) in Λ_1 (resp. Λ_2) are simply the values $\alpha(\mu)$, where α denotes an arbitrary root of M_1 in Λ_1 (resp. M_2 in Λ_2). But then $(\mu, \mu)_1 = \mathrm{tr}\,(\mathrm{Ad}\,(\mu)\,\mathrm{Ad}\,(\mu)) = \sum (\alpha(\mu))^2$, the sum being extended over all roots of M_1 in Λ_1, so that by Lemma II.3.4 $(\mu, \mu)_1$ must be positive. In the same way $(\mu, \mu)_2$ must be positive, and the second assertion follows. Q.E.D.

Now we give the proof of Lemma 9.

Proof of Lemma 9. We proceed case-by-case through the list of simple Lie algebras given by Theorem II.3.36, using the models for these Lie algebras established in Section II.4 as convenient. In the following proof, in discussing any particular complex simple Lie algebra, we shall use the notations introduced in the discussion of that Lie algebra in section II.4.

A Lie algebra of type A_{n-1} may be taken to be the Lie algebra of all $n \times n$ complex matrices λ of trace zero. The reader will easily verify that the natural bilinear form in A_{n-1} is $2n\,\mathrm{tr}\,(\lambda\mu)$. This Lie algebra admits the automorphism $\lambda \to -\lambda'$, where λ' is the transpose matrix of the matrix λ, and hence the conjugation $C: \lambda \to -\bar{\lambda}'$. Since $(\lambda, C\lambda) = -2n\,\mathrm{tr}\,(\lambda\bar{\lambda}')$, we see that Lemma 9 is verified for algebras of this type.

A Lie algebra of type B_m, C_m or D_m may be taken to be the Lie algebra of all $p \times p$ complex matrices λ of trace zero satisfying $\langle\langle \lambda x, y \rangle\rangle = -\langle\langle x, \lambda y \rangle\rangle$, where we put $p = 2m + 1$ if we deal with B_m, $p = 2m$ if we deal with C_m or D_m, and

$$\langle\langle x, y \rangle\rangle = x_1 y_p + \cdots + x_p y_1$$

if we deal with B_m or D_m,

$$\langle\langle x, y \rangle\rangle = x_1 y_p + \cdots + x_m y_{m+1} - x_{m+1} y_m - \cdots - x_p y_1$$

if we deal with C_m. In all cases, the set of diagonal matrices is a Cartan subalgebra. By Lemma 10, the natural bilinear form is $a \cdot \operatorname{tr}(\lambda\mu)$, where a is a constant. Putting $\lambda_0 = e_{11} - e_{pp}$, and noting that (λ_0, λ_0) is the sum over all roots α of $(\alpha(\lambda_0))^2$, while by section II.4, $\alpha(\lambda_0)$ is always real, we see that $a > 0$. Let J be the matrix of the transformation $[x_1, \dots, x_p] \to [x_p, \dots, x_1]$ if we deal with B_m or D_m, and the matrix of the transformation $[x_1, \dots, x_{2m}] \to [x_{2m}, \dots, x_{m+1}, -x_m, \dots, -x_1]$ if we deal with C_m. Then it is easily verified that $\langle\langle \lambda x, y\rangle\rangle = -\langle\langle x, \lambda y\rangle\rangle$ is equivalent to $\lambda' = -J\lambda J^{-1}$, where λ' is the transpose of the matrix λ. Thus the conjugation $C: \lambda \to J\bar{\lambda}J^{-1} = -\bar{\lambda}'$ verifies Lemma 9 for algebras of these three types.

A model for the Lie algebra of type G_2 is given as a direct sum $A \oplus V \oplus V'$ of the Lie algebra A of type A_2, consisting of all 3×3 matrices of trace zero, a 3-dimensional vector space V, and the three dimensional "dual" vector space V' of all linear functionals on V. The set M of all diagonal matrices in A is a Cartan subalgebra. Using Lemma 10, and the description of the roots and root vectors given in section II.4, we see that the natural bilinear form of G_2 satisfies $(\lambda + v + v', \lambda + v + v') = a \operatorname{tr}(\lambda)^2 + (b_1 v_1 v_1' + b_2 v_2 v_2' + b_3 v_3 v_3')$; but then plainly $b_1 = b_2 = b_3$ by symmetry, so that the natural bilinear form is $a \cdot \operatorname{tr}(\lambda^2) + bv'(v)$. Using section II.4 (cf. especially formula (7) of that section) and the second part of Lemma 10 we find $a > 0$. Then, using the identity $(\operatorname{Ad}(v')\,\lambda, v) = -(\lambda, \operatorname{Ad}(v')\,v)$, i.e. $b(\lambda' v')\,(v) = -a \cdot \operatorname{tr}(\lambda(v' \cdot v))$, we find $3b = -a$, so that $b < 0$ (cf. the last sentence preceeding formula (8) of section 2.4.). Since the mapping $T: v \to u', v' \to u$ defined by $[u_1', u_2', u_3'] = (2/27)^{-1/3} [v_1, v_2, v_3]$, $[u_1, u_2, u_3] = (2/27)^{1/3} [v_1', v_2', v_3']$ satisfies $v'(v) = (Tv)(Tv')$, $\langle Tx, Ty, Tz\rangle = \langle x, y, z\rangle$, and $\langle Tx', Ty', Tz'\rangle = \langle x, y, z\rangle$, an examination of the detailed definition of the Lie product in our Lie algebra, as given in section II.4, will show that G_2 admits the automorphism $\lambda \to -\lambda'$, $v \to Tv$, $v' \to Tv'$, and hence the conjugation $C: \lambda \to -\bar{\lambda}'$, $v \to \bar{T}\bar{v}$, $v' \to \bar{T}\bar{v}'$, which from the above verifies Lemma 9 for the case of an algebra of type G_2.

A model for the Lie algebra of type E_8 is given as a direct sum $A \oplus V \oplus V'$ of the Lie algebra A of type A_8, consisting of all 9×9 matrices of trace zero, the vector space V of all tensors v_{ijk} of three indices, skew-symmetric in the three indices, and the "dual" vector space V'. Arguing just as in the preceding paragraph, the reader will verify that the natural bilinear form of E_8 satisfies $(\lambda + v + v', \lambda + v + v') = a \cdot \operatorname{tr}(\lambda^2) + bv'(v)$ with $a > 0$ and $b < 0$, and that E_8 admits the automorphism $\lambda \to -\lambda'$, $v \to Tv = u'$, $v' \to Tv' = u$, where $(u')^{ijk} = 2^{-5/3}3^{-7/3}7^{-1/3}v_{ijk}$, u_{ijk}

$= 2^{5/3}3^{7/3}7^{1/3}(v')^{ijk}$. Thus the conjugation $C : \lambda \to -\bar{\lambda}'$, $v \to \bar{T}\bar{v}$, $v' \to \bar{T}\bar{v}'$ verifies Lemma 9 for an algebra of type E_8.

A model for the Lie algebra of type E_7 is given by the subalgebra of the above Lie algebra of type E_8 spanned by all those vectors μ_α and $\lambda_{\pm\alpha}$ for which the root α is linearly dependent on the roots $\omega_2 - \omega_3, \ldots, \omega_7 - \omega_8$, $\omega_6 + \omega_7 + \omega_8$ (cf. the relevant passages in section II.4). Using the description of the vectors μ_α and λ_α given in section II.4, the reader will verify readily the conjugation C maps μ_α to $\mu_{-\alpha}$, and maps λ_α to a multiple of $\lambda_{-\alpha}$. Thus C maps the subalgebra E_7 of E_8 into itself, and can be regarded as a conjugation in E_7. By Lemma 10, the natural bilinear form in E_7 is a constant b times the restriction of the natural bilinear form of E_8 to E_7; since by Lemma 10 we have $(\mu_{\omega_2-\omega_3}, \mu_{\omega_2-\omega_3}) > 0$ for both the natural bilinear form of E_8 and that of E_7, we must have $b > 0$. Therefore the restriction to E_7 of the conjugation C verifies Lemma 9 for an algebra of type E_7.

A model for the Lie algebra $\hat{\Lambda}^{(1)}$ of type E_6 is given by the subalgebra of the above Lie algebra of type E_8 spanned by all those vectors μ_α and $\lambda_{\pm\alpha}$ for which the root α is linearly dependent on the roots $\omega_3 - \omega_4, \ldots$, $\omega_7 - \omega_8, \omega_6 + \omega_7 + \omega_8$ (cf. the relevant passages in section II.4). Arguing exactly as in the preceding paragraph, we see that the conjugation C in E_8 constructed above maps E_6 into itself, and that the restriction to E_6 of this conjugation verifies Lemma 9 for an algebra of type E_6.

According to the relevant passages of section II.4, the simple complex Lie algebra $\hat{\Lambda} = \Lambda \oplus V \oplus V'$ of type E_8 admits as a Cartan subalgebra the set $\hat{M}$ of all diagonal matrices in Λ. As we have noted, the conjugation C in this simple Lie algebra Λ has the property $C\hat{\Lambda}_\alpha = \hat{\Lambda}_{-\alpha}$, where α denotes any root of $\hat{M}$ in $\hat{\Lambda}$, and $\hat{\Lambda}_\alpha$ denotes the corresponding root space. It is then evident that the restriction of C to the subalgebras of types E_7 and E_6 discussed above has this same property. Let $0 \neq \lambda_\alpha \in \hat{\Lambda}_\alpha$. Since $(\lambda_\alpha, C\lambda_\alpha) < 0$, we may, replacing λ_α by a multiple of itself if necessary, suppose that $(\lambda_\alpha, C\lambda_\alpha) = -1$. This shows that we may choose vectors $\lambda_\alpha \in \hat{\Lambda}_\alpha, \alpha \neq 0$, in such a way that $(\lambda_\alpha, \lambda_{-\alpha}) = 1$, and $C\lambda_\alpha = -\lambda_{-\alpha}$.

Let the roots $\omega_3 - \omega_4, \ldots, \omega_7 - \omega_8, \omega_6 + \omega_7 + \omega_8$ of $\hat{M}$ in $\hat{\Lambda}$ be called $\alpha_1, \ldots, \alpha_5, \alpha'$ for notational convenience. By Corollary 3.22, there is an automorphism Φ (unique, by Corollary 3.19) of the Lie algebra $\hat{\Lambda}^{(1)}$ of type E_6 such that $\Phi : \lambda_{\pm\alpha_1}, \lambda_{\pm\alpha_2}, \lambda_{\pm\alpha_3}, \lambda_{\pm\alpha_4}, \lambda_{\pm\alpha_5}, \lambda_{\pm\alpha'} = \lambda_{\pm\alpha_5}, \lambda_{\pm\alpha_4}, \lambda_{\pm\alpha_3}, \lambda_{\pm\alpha_2}, \lambda_{\pm\alpha_1}, \lambda_{\pm\alpha'}$. By the relevant passages of section II.4, the subalgebra $\tilde{\Lambda} = \{\hat{\lambda} \in \hat{\Lambda}^{(1)} | \Phi\hat{\lambda} = \hat{\lambda}\}$ of $\hat{\Lambda}^{(1)}$ is of type F_4. But it is clear from the definition of Φ that $C\Phi C = \Phi$; thus $C\tilde{\Lambda} \subseteq \tilde{\Lambda}$, and the restriction of C to $\tilde{\Lambda}$ is a conjugation in $\tilde{\Lambda}$. Noting the description of a Cartan subalgebra $\hat{M}$ of $\tilde{\Lambda}$

and of the roots of $\tilde{M}$ in $\tilde{\Lambda}$ as given in section II.4, we see that the second part of Lemma 10 is applicable, so that the natural bilinear form in $\tilde{\Lambda}$ is a positive multiple of the restriction to $\tilde{\Lambda}$ of the natural bilinear form in $\tilde{\Lambda}$. But then it is plain that the restriction of C to $\tilde{\Lambda}$ is a positive conjugation in $\tilde{\Lambda}$ so that Lemma 9 is verified in all cases. Q.E.D.

11. Corollary. Let $\hat{\Lambda}$ be a semi-simple complex Lie algebra. Then there exists a positive conjugation C in $\hat{\Lambda}$.

Proof. This follows trivially from Lemma 9 and from the decomposition, given by Theorem 2.17 of $\hat{\Lambda} = \Lambda_1 + \Lambda_2 + \cdots + \Lambda_n$ into the direct sum of ideals Λ_i, each of which is a simple Lie algebra. Q.E.D.

The next lemma recapitulates some useful information on the relation of the group of inner automorphisms of $\hat{\Lambda}$ to the group of all automorphisms of $\hat{\Lambda}$; the one which follows it enables us to show in various situations that a given automorphism is inner.

12. Lemma. Let $\hat{\Lambda}$ be a complex semi-simple Lie algebra. Let A be its group of automorphisms, and A_0 the subgroup consisting of all inner automorphisms. Then A_0 is a normal subgroup of A; the factor group A/A_0 is finite. If $\hat{\Lambda}$ is simple, the order of the factor group A/A_0 is at most equal to the number of symmetries of the Dynkin diagram of $\hat{\Lambda}$.

Proof. A_0 is the subgroup of A generated by all the automorphisms $\exp(\mathrm{Ad}(\lambda))$, $\lambda \in \hat{\Lambda}$. If $\psi \in A$, we have $\psi\, \mathrm{Ad}(\lambda)\, \psi^{-1} = \mathrm{Ad}(\psi(\lambda))$, thus $\psi \exp(\mathrm{Ad}(\lambda))\, \psi^{-1} = \exp(\mathrm{Ad}(\psi(\lambda))$, proving that A_0 is a normal subgroup of A. Let $\tilde{M}$ be a Cartan subalgebra of $\hat{\Lambda}$ and $\{\alpha\}$ the roots of $\tilde{M}$ in Λ. Then, by Theorem II.5.1 and the remark which follows its statement, every $\psi \in A$ can be written as $\psi'\phi$, where $\phi \in A_0$, and where $\psi'(\tilde{M}) = \tilde{M}$. If λ is a root vector belonging to the root α, then $\mathrm{Ad}(\mu)\, \psi'(\lambda) = \psi'\, \mathrm{Ad}((\psi')^{-1}\mu)\, \lambda = \alpha((\psi')^{-1}\mu)\, \psi'(\lambda)$, so that $\psi'(\lambda)$ is also a root vector, belonging to the root ψ'_α defined by $\alpha((\psi')^{-1}\mu) = (\psi'_\alpha)(\mu)$. It follows that ψ' maps the root space Λ_α onto the root space Λ'_α, and hence maps the vector μ_α (cf. Corollary II.3.8) onto the vector μ'_α. If ψ'' is a second automorphism mapping $\tilde{M}$ onto itself, and $\psi''_\alpha = \psi'_\alpha$ for all $\alpha \in \{\alpha\}$, then $\psi''(\psi')^{-1}$ is inner by Lemma II.5.11. Thus the order of A/A_0 is at most equal to the number of permutations of the roots $\{\alpha\}$.

Next, let $\hat{\Lambda}$ be simple, and let Π be a simple set of roots of $\tilde{M}$ in $\hat{\Lambda}$. Then, by Theorem II.5.1 and the remark which follows its statement, and by Theorem II.5.4, every $\psi \in A$ can be written as $\psi = \psi'\phi$ where $\phi \in A_0$ and $\psi'_\alpha \in \Pi$ if $\alpha \in \Pi$. It is clear from Definition II.3.24 that ψ'_α is a simple, double, or triple neighbor of ψ'_β if and only if α is a simple, double, or triple neighbor of β. Thus, by Definition II.3.35, $\alpha \to \psi'_\alpha$, $\alpha \in \Pi$, gives a symmetry of the Dynkin diagram. If ψ'' is a second automorphism mapping M into itself and such that $\psi'_\alpha \in \Pi$ if $\alpha \in \Pi$, and if $\psi''_\alpha = \psi'_\alpha$ for all $\alpha \in \Pi$, then $\psi''(\psi')^{-1}$ is inner by Lemma 5.11 and Corollary 3.10. Thus the order of A/A_0 is at most equal to the number of symmetries of the Dynkin diagram of Λ. Q.E.D.

13. Lemma. Let ψ be a linear transformation in a semi-simple Lie algebra $\hat{\Lambda}$ such that $\psi[\lambda, \mu] = [\psi\lambda, \mu] + [\lambda, \psi\mu]$, $\lambda, \mu \in \hat{\Lambda}$. Then there exists an element $\lambda_0 \in \hat{\Lambda}$ such that $\psi = \text{Ad}(\lambda_0)$.

Proof: The expression $\text{tr}(\text{Ad}(\lambda)\,\psi)$ defines a linear functional of λ, thus, by Theorem II.2.3 and Lemma 0.5 there exists a $\lambda_0 \in \hat{\Lambda}$ such that $\text{tr}(\text{Ad}(\lambda)\,\psi) = \text{tr}(\text{Ad}(\lambda)\,\text{Ad}(\lambda_0))$, $\lambda \in \hat{\Lambda}$. Thus, if we put $\phi = \psi - \text{Ad}(\lambda_0)$, we have $\phi[\lambda, \mu] = [\phi\lambda, \mu] + [\lambda, \phi\mu]$ and $\text{tr}(\text{Ad}(\lambda)\,\phi) = 0$ for all $\lambda, \mu \in \hat{\Lambda}$.

Therefore if $\lambda, \mu \in \hat{\Lambda}$ we have

$$\begin{aligned}\text{tr}(\text{Ad}(\phi\lambda)\,\text{Ad}(\mu)) &= \text{tr}((\phi\,\text{Ad}(\lambda) - \text{Ad}(\lambda)\,\phi)\,\text{Ad}(\mu)) \\ &= \text{tr}(\phi\,\text{Ad}(\lambda)\,\text{Ad}(\mu) - \text{Ad}(\mu)\,\text{Ad}(\lambda)\,\phi) \\ &= \text{tr}(\phi\,\text{Ad}([\lambda, \mu])) = 0.\end{aligned}$$

It follows by Theorem II.2.3 that $\phi\lambda = 0$ for $\lambda \in \hat{\Lambda}$, and thus $\phi = 0$ so that $\psi = \text{Ad}(\lambda_0)$. Q.E.D.

Using Lemmas 9 and 13, we can carry over the following very important theorem of linear algebra to the group of automorphisms of a general Lie algebra.

14. Lemma. ("Canonical Factorization"): Let ϕ be an automorphism of the semi-simple complex Lie algebra $\hat{\Lambda}$ .Let C be a positive conjugation in $\hat{\Lambda}$. Then we may factor ϕ as $\phi = up$, where u and p are automorphisms of $\hat{\Lambda}$, $u = u_c$, $pp_c = I$, p is inner and p commutes with every automorphism ψ such that $\phi_c^{-1}\phi\psi = \psi\phi_c^{-1}\phi$.

Proof. For each linear transformation T of $\hat{\Lambda}$ into itself, let T^* be the "adjoint" defined by $\langle T\lambda, \mu\rangle = \langle\lambda, T^*\mu\rangle_c$, $\lambda, \mu \in \hat{\Lambda}$. We have $(ST)^* = T^*S^*$; thus, if T has an inverse, so has T^*, and so has $T^*T = T_1$. We have $T^*T = T_1$ and $\langle T_1\lambda, \lambda\rangle_c = \langle T\lambda, T\lambda\rangle_c \geqslant 0$ for $\lambda \in \hat{\Lambda}$; thus, by Lemma 0.26, there exists a unique linear transformation P such that $\langle P\lambda, \lambda\rangle_c \geqslant 0$ for $\lambda \in \hat{\Lambda}$, and $P^2 = T_1$. Plainly, P has an inverse. Put $U = TP^{-1}$; then $U^*U = P^{-1}T^*TP^{-1} = P^{-1}P^2P^{-1} = I$. This shows that T may be factored as $T = UP$ where $UU^* = I$ and $\langle P\lambda, \lambda\rangle_c \geqslant 0$. Since $T = UP$, $P = P^*$ and $U^*U = I$ together imply that $T^*T = P^2$, we see that P is unique; thus U is also unique. Moreover, by Lemma 0.26, P commutes with every linear transformation commuting with T^*T.

Apply the above factorization to ϕ, to get $\phi = up$, where $uu^* = 1$, $p = p^*$, and $\langle p\lambda, \lambda\rangle_c \geqslant 0$ for $\lambda \in \hat{\Lambda}$. Since ϕ is an automorphism of $\hat{\Lambda}$, we have $(\phi\lambda, \phi\mu) = (\lambda, \mu)$ for $\lambda, \mu \in \hat{\Lambda}$, and thus

$$\begin{aligned}\langle\phi\lambda, \mu\rangle_c &= -(\phi\lambda, C\mu) = -(\lambda, \phi^{-1}C\mu) = -(\lambda, C^2\phi^{-1}C\mu)\\ &= \langle\lambda, \phi_c^{-1}u\rangle_c;\end{aligned}$$

thus $\phi^* = \phi_c^{-1}$, or $(\phi^*)_c^{-1} = \phi$. Hence we have $\phi = (u^*)_c^{-1}(p^*)_c^{-1} = u\,p_c^{-1}$. The reader will verify easily that $(u_c)^*\,u_c = I$, $p_c^* = p_c$, and $\langle p_c\lambda, \lambda\rangle_c \geqslant 0$ for $\lambda \in \Lambda$. Thus, from the uniqueness of the factorization $\phi = up$ as established above, we find that $u = u_c$, $p = p_c^{-1}$. Moreover, $p^2 = \phi^*\phi = \phi_c^{-1}\phi$ is an automorphism. We have only to show that p is an inner automorphism, and our lemma will be proved.

This we do as follows. Since $p = p^*$, it follows by Lemma 0.23 that $\hat{\Lambda}$ decomposes as a direct sum $\hat{\Lambda} = \Lambda(\alpha_1) \oplus \cdots \oplus \Lambda(\alpha_k)$ of eigenspaces belonging to the various eigenvalues $\alpha_1, \ldots, \alpha_k$ of p, all of which are positive by Lemma 0.25. If $\lambda \in \Lambda(\alpha_j)$ and $\lambda' \in \Lambda(\alpha_{j'})$, then $p^2[\lambda, \lambda'] = [p^2\lambda, p^2\lambda'] = (\alpha_j\alpha_{j'})^2\,[\lambda, \lambda']$. Since all the eigenvalues of p are positive, it follows that $p[\lambda, \lambda'] = \alpha_j\alpha_{j'}[\lambda, \lambda']$.

Define a mapping ψ by putting $\psi\lambda = \log(\alpha_j)\,\lambda$ for $\lambda \in \Lambda(\alpha_j)$. Then $\psi[\lambda, \lambda'] = \{\log(\alpha_j) + \log(\alpha_{j'})\}\,[\lambda, \lambda'] = [\psi\lambda, \lambda'] + [\lambda, \psi\lambda']$ for $\lambda \in \Lambda(\alpha_j)$ and $\lambda' \in \Lambda(\alpha_{j'})$, which makes it plain that $\psi[\lambda, \lambda'] = [\psi\lambda, \lambda'] + [\lambda, \psi\lambda']$ for all $\lambda, \lambda' \in \hat{\Lambda}$. Thus, by Lemma 13, there exists a $\mu \in \hat{\Lambda}$ such that $\psi = \mathrm{Ad}\,(\mu)$. We have $\exp(\mathrm{Ad}\,(\mu))\,\lambda = \exp(\psi)\,\lambda = \exp(\log(\alpha_j))\,\lambda = \alpha_j\lambda$ for $\lambda \in \Lambda(\alpha_j)$; thus $\exp(\mathrm{Ad}\,(\mu)) = p$, and p is an inner automorphism. Q.E.D.

15. Corollary. The automorphisms u and p of the preceding lemma are unique. We have $\langle p\lambda, \lambda\rangle_c \geqslant 0$ for all $\lambda \in \hat{\Lambda}$. There exists a unique inner automorphism $p^{1/2}$ such that $(p^{1/2})^2 = p$, and $\langle p^{1/2}\lambda, \lambda\rangle_c \geqslant 0$ for all $\lambda \in \hat{\Lambda}$; moreover, p and $p^{1/2}$ commute with every transformation which commutes with p^2.

Proof. If we let μ be as in the preceding proof, and put $p_0 = \exp(\frac{1}{2}\,\mathrm{Ad}\,\mu)$, then p_0 is an inner automorphism, and plainly $p_0^2 = p$. We have $p_0\lambda = \alpha_j^{1/2}\lambda$ for $\lambda \in \Lambda(\alpha_j)$. By Lemma 0.23 we have $\langle\Lambda(\alpha_j), \Lambda(\alpha_k)\rangle = 0$ if $\alpha_j \neq \alpha_k$, from which it follows trivially that $\langle p_0\lambda, \lambda\rangle \geqslant 0$ for all $\lambda \in \hat{\Lambda}$. This shows the existence of $p_0 = p^{1/2}$; the uniqueness and our final assertion follows from Lemma 0.26. Q.E.D.

16. Corollary. Every conjugation $C_1 = \phi C$ in $\hat{\Lambda}$ is equivalent to a conjugation of the form uC, where the automorphism u satisfies $u_c = u$, $u^2 = I$, and where the automorphism $u^{-1}\phi$ is inner.

Proof. Factor $\phi = up$ as in Lemma 14; then $upCupC = upu_cp_c = upup^{-1} = I$. Hence $u^{-1} = pup^{-1}$, which implies $u = pu^{-1}p^{-1}$, and hence $u = p^2up^{-2}$. But then u commutes with p by the preceding corollary; so that $u^{-1} = u$, i.e., $u^2 = I$. The conjugation C_1 is equivalent to $p^{1/2}uCp^{1/2}$. It is easily verified that $\langle C(p^{1/2})^{-1} C\lambda, \lambda\rangle \geqslant 0$ for all $\lambda \in \hat{\Lambda}$, and of course $(C(p^{1/2})^{-1} C)^2 = Cp^{-1}C = p$; thus $C(p^{1/2})^{-1} C = p^{1/2}$, and $Cp^{1/2}C = (p^{1/2})^{-1}$. Therefore C_1 is equivalent to $(p^{1/2})\,up^{-1/2}C = uC$; since $u^{-1}\phi = p$ is inner the present corollary is proved. Q.E.D.

17. Corollary. Let ϕ and ψ be two automorphisms in $\hat{\Lambda}$, and suppose that $\phi_c = \phi$, $\psi_c = \psi$. Suppose that there exists an automorphism η such that $\eta\phi\eta^{-1} = \psi$. Then there exists an automorphism u such that $u\phi u^{-1} = \psi$, and such that $u = u_c$.

Proof: We have $\eta\phi\eta^{-1} = \psi$, and thus $\eta_c\phi\eta_c^{-1} = \psi$. Therefore $\eta_c^{-1}\eta$ commutes with ϕ; thus, using the Lemma 14, we may write $\eta = up$, where $u = u_c$ is an automorphism, and where p commutes with ϕ. But then $u\phi u^{-1} = \psi$. Q.E.D.

18. Corollary. Let uC and $u'C$ be two conjugations in $\hat{\Lambda}$; we suppose that $u_c = u$, $u'_c = u'$, $u^2 = I$, $(u')^2 = I$. Suppose also that there exists an automorphism ϕ of $\hat{\Lambda}$ such that $\phi u\phi^{-1} = u'$. Then the conjugations uC and $u'C$ are equivalent.

Proof. By Corollary 17, we can suppose $\phi_c = \phi$; but then $\phi uC\phi^{-1} = \phi u\phi^{-1}C = u'C$. Q.E.D.

To verify that two real simple Lie algebras derived from a single complex simple Lie algebra are not isomorphic, directly by the method which Lemma 6 suggests, will not always be convenient. The following definition establishes an invariant which will in many cases yield easier proofs of non-isomorphism.

19. Definition. Let Λ_1 be a real semi-simple Lie algebra, and let (λ, μ) be its natural bilinear form. The signature of (λ, μ) will be called the signature of Λ_1.

Note what is obvious: That isomorphic real semi-simple algebras have equal signatures. The following simple lemma, which is related to Lemma 10, will facilitate our computations of the signature in a number of cases.

20. Lemma. Let Λ be a complex Lie algebra, C_1 a conjugation in $\hat{\Lambda}$, and Λ the real Lie algebra defined by the condition $C_1\lambda = \lambda$. Then the natural bilinear form (λ, μ) of Λ is related to the natural bilinear form $(\lambda, \mu)_1$ of Λ by the equation $(\lambda, \mu) = (\lambda, \mu)_1$, $\lambda, \mu \in \Lambda$.

Proof. Let $v_1, \ldots, v_l$ be a basis for the real vector space Λ; we first show that $v_1, \ldots, v_l$ is also a basis for the complex vector space $\hat{\Lambda}$. Indeed, if $\eta \in \hat{\Lambda}$, we may write $\eta = \frac{1}{2}(\eta + C\eta) + i\frac{1}{2i}(\eta - C\eta)$. Since $\frac{1}{2}(\eta + C\eta)$ and $\frac{1}{2i}(\eta - C\eta)$ both belong to Λ, it is clear that $v_1, \ldots, v_l$ span $\hat{\Lambda}$. On the other hand, if $z_1v_1 + \cdots + z_lv_l = 0$, then it is clear on applying C_1, adding, and subtracting, that $Re(z_1)v_1 + \cdots + Re(z_l)v_l = 0 = Im(z_1)v_1 + \cdots + Im(z_l)v_l = 0$. Thus $z_1 = \cdots = z_l = 0$, showing that $v_1, \ldots, v_l$ is a basis for $\hat{\Lambda}$. Therefore, if $\lambda, \mu \in \Lambda$, the matrix of the linear transformation Ad (λ) Ad (μ) in the basis $v_1, \ldots, v_l$ is the same whether this transformation be regarded as acting in the vector space Λ or in the vector space $\hat{\Lambda}$; and from this remark our Lemma follows immediately. Q.E.D.

It is now easy for us to give a complete analysis of the conjugations in a simple complex algebra of any of the four "classical" types A_m, B_m, C_m, D_m, and, without further ado, we proceed to this task. We shall work case-by-case through each of the four classical types, making reference to the specific models established in section II.4 as convenient. We first give the

A. Analysis of the conjugations in the Lie algebra of type A_{n-1}

As we saw in section II.4, this Lie algebra $\hat{\Lambda}$ may be taken to be the Lie algebra of all $n \times n$ complex matrices of trace zero. By Definition II.5.3. and Lemma I.5.3, every inner automorphism of $\hat{\Lambda}$ has the form $\lambda \to T\lambda T^{-1}$, where T is non-singular. By Lemma 12, the number of cosets of the group A_0 of inner automorphisms of $\hat{\Lambda}$ in the group A of all automorphisms of $\hat{\Lambda}$ is at most 2 in general, and at most 1 in the degenerate case $n = 2$. Now, $\hat{\Lambda}$ admits the automorphism $\lambda \to -\lambda'$, where λ' is the transpose of the matrix λ. Except when $n = 2$, it is plain that there can exist no matrix S such that $S\lambda S^{-1} = -\lambda'$ for all $\lambda \in \hat{\Lambda}$, or even for all diagonal $\lambda \in \hat{\Lambda}$; thus the automorphism $\lambda \to -\lambda'$ is not inner, and $(n > 2)$ A/A_0 is of order 2. Every automorphism is then either of the "inner" form $\lambda \to T\lambda T^{-1}$ or of the "outer" form $\lambda \to -T\lambda' T^{-1}$, where T is non-singular. It is plain that in both these cases the automorphism determines the matrix T up to a constant factor.

We saw in proving Lemma 9 that the natural bilinear form of $\hat{\Lambda}$ is $(\lambda, \mu) = 2n \operatorname{tr} (\lambda\mu)$, and that $C : \lambda \to -\bar{\lambda}'$ is a positive conjugation. By Corollary 16, every conjugation in $\hat{\Lambda}$ is equivalent to a conjugation of the form ϕC, where the automorphism ϕ satisfies $\phi^2 = I$ and $\phi \; = \phi$. Thus ϕ is either of the form $\lambda \to T\lambda T^{-1}$ with $T^2 = aI$ and $T\bar{T}' = bI$, or of the form $\lambda \to -T\lambda' T^{-1}$ with $T(T')^{-1} = aI$ and $T\bar{T}' = bI$. We examine these two cases separately.

In the first case, replacing T by a constant multiple of itself as necessary, we will have $T^2 = I$, and we may suppose in addition that the dimension of the eigenspace V_{+1} of T is at least as great as the dimension of the eigenspace V_{-1}. If S is any other transformation satisfying $S^2 = I$ and $S\bar{S}' = b'I$, and if $S = VTV^{-1}$ for some non-singular V, it follows by Corollary 18 that the conjugation ϕC is equivalent to the conjugation $\lambda \to SC(\lambda)\, S^{-1}$. But then, if $n - k$ is the dimension of V_+, we may clearly take S to be the diagonal matrix I^{+}_{n-k} of $n - k$ 1's followed by k $-$ 1's; here $k \leqslant n/2$.

In the second case $T\bar{T}' = bI$, and taking traces we see that $b > 0$; replacing T by a suitable positive multiple of itself, we may take $b = 1$. Then $\bar{T}' = T^{-1}$, so $T' = \bar{T}^{-1}$, $(T')^{-1} = \bar{T}$, and $T\bar{T} = \bar{T}T = aI$. Hence a is real, and taking determinants of both sides of the equation $T\bar{T} = aI$, we find that $a = \pm 1$, with $a = -1$ impossible if n is odd. Our conjugation ϕC is then $\lambda \to T\bar{\lambda}T^{-1}$. By Lemma 0.23, there exists a complete basis for the space of all vectors, consisting of eigenvectors of the matrix T. We analyze the two cases $a = +1$ and $a = -1$ separately.

(+) $a = +1$. Let the eigenspace of T belonging to the eigenvalue z be called V_z. Then, since $T\bar{T} = 1$, we have $\bar{v} \in V_{\bar{z}^{-1}}$ if $v \in V_z$. For each z in the spectrum of T choose a complex number $\zeta(z)$, and define N by $Nv = \zeta(z)\, v$ if $v \in V_z$. Then $\bar{N}^{-1}\, Nv = \overline{\zeta(\bar{z}^{-1})^{-1}}\, \zeta(z)\, v$ for $v \in V_z$. If we choose $\zeta(z)$ so that $\overline{\zeta(\bar{z}^{-1})^{-1}}\, \zeta(z) = z$ for each eigenvalue z of T, we find $\bar{N}^{-1}N = T$; thus the conjugation $\lambda \to T\bar{\lambda}T^{-1}$ is equivalent to $\lambda \to \bar{\lambda}$.

(−) $a = -1$. Using the same notations as immediately above, we find $v \in V_{-\bar{z}^{-1}}$ if $v \in V_z$. Defining N in terms of $\zeta(z)$ as immediately above, we find $\bar{N}^{-1}TN = \overline{\zeta(-\bar{z}^{-1})^{-1}}\, \zeta(z) z$. If we divide the spectrum of T into pairs $z, -z'^{-1}$, and choose $\zeta(z)$ so that $\overline{\zeta(-\bar{z}^{-1})^{-1}}\, \zeta(z)\, z = 1$, we see that the conjugation $\lambda \to T\bar{\lambda}T^{-1}$ is equivalent to a conjugation $\lambda \to T_1\bar{\lambda}T_1^{-1}$, where T_1 is a matrix which has only the two eigenvalues ± 1, which is such that there exists a basis consisting of eigenvectors of T_1, and which is such that $T_1 v' = -v'$ if $T_1 v = v$. It is plain that any two such transformations are equivalent by a *real* change of basis; thus the conjugation $\lambda \to T_1\bar{\lambda}T_1^{-1}$

is equivalent to the conjugation $\lambda \to \mathscr{Y}_{2k}\bar{\lambda} \cdot \mathscr{Y}_{2k}$, where $2k = n$ and where $\mathscr{Y}_{2k}$ is the matrix

$$\mathscr{Y}_{2k} = i \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & & \\ & \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & \\ & & \ddots \end{pmatrix}. \tag{1}$$

The conjugations $\lambda \to -I^{+}_{n-k}\,\bar{\lambda}' I^{+}_{n-k}$ which arose above under our "first case" define the real Lie algebras $HA^{(k)}_{n-1}$ of all matrices λ of trace zero satisfying $\lambda = -I^{+}_{n-k}\,\bar{\lambda}' I^{+}_{n-k}$; $HA^{(k)}_{n-1}$ may be regarded as the set of all matrices λ of trace zero satisfying $\{\lambda u, v\} = \{v, \lambda u\}$, where $\{u, v\}$ is the hermitian form

$$\{u, v\} = u_1\bar{v}_1 + \cdots + u_{n-k}\bar{v}_{n-k} - u_{n-k+1}\bar{v}_{n-k+1} - \cdots - u_n\bar{v}_n.$$

The Lie algebra of all real $n \times n$ matrices of trace zero we call RA_{n-1}; the Lie algebra of all $n \times n$ matrices of trace zero satisfying $\bar{\lambda} = \mathscr{Y}_{2k}\lambda\mathscr{Y}_{2k}$ we call QA_{n-1} in case $n = 2k$. This latter algebra may be regarded as the Lie algebra of all $k \times k$ matrices of trace zero whose elements are elements of the 4-dimensional real division algebra determined by $\bar{\lambda} = \mathscr{Y}_2\lambda\mathscr{Y}_2$, i.e., the division algebra of 2×2 matrices

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix},$$

which is, of course, the algebra of quaternions.

We shall now verify that no two of the simple real Lie algebras we have just listed are isomorphic. This is most conveniently done by calculating the signature of the algebras. The natural bilinear form of the complex Lie algebra A_{n-1} is $2n \operatorname{tr}(\lambda\lambda')$, and therefore, if the field of scalars be restricted from the complex field to the real field, $4n \operatorname{tr}(\lambda\lambda')$. By Lemma 10, $\operatorname{tr}(\lambda\lambda')$ is then a positive multiple of the natural bilinear form of each of the real algebras we have constructed.

The signature of RA_{n-1} is thus seen to be the number of parameters on the diagonal of a matrix $\lambda \in RA_{n-1}$, plus half the number of the remaining parameters; thus $(1 + \cdots + (n-1)) + (n-1) = \dfrac{n(n-1)}{2} + n - 1 = \dfrac{(n+2)(n-1)}{2}$.

We obtain the signature of the algebra QA_{n-1}. Since

$$\operatorname{tr}\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}\begin{pmatrix} a_1 & b_1 \\ -\bar{b}_1 & \bar{a}_1 \end{pmatrix} = aa_1 + \bar{a}_1\bar{a} - b\bar{b}_1 - b_1\bar{b}$$

may be written as the difference of four squares and four squares, and since

$$\operatorname{tr}\left(\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}^2\right) = a^2 + \bar{a}^2 - 2\,|b|^2$$

is the difference of a square and three squares, the signature of QA_{n-1} is $k - 1 + 4(1 + \cdots + k - 1) = 2k(k-1) + k - 1 = (2k+1)(k-1)$ $= \dfrac{(n+1)(n-2)}{2}$. Thus RA_{n-1} and QA_{n-1} are never isomorphic.

Calling the automorphisms $\lambda \to T\lambda T^{-1}$ *inner*, and the automorphisms $\lambda \to T\lambda' T^{-1}$ *outer*, we see that the product of two outer automorphisms is inner, etc. Therefore, a conjugation of the form $\lambda \to T\bar{\lambda}T^{-1}$ can never be equivalent to a conjugation of the form $\lambda \to -T\bar{\lambda}'T^{-1}$ except in the degenerate case $n = 2$.

The algebra $HA^{(k)}_{n-1}$ consists of all $n \times n$ matrices of trace zero which may be partitioned as $\begin{pmatrix} ih_1 & m \\ m^* & ih_2 \end{pmatrix}$, where we write T^* for the conjugate transpose of the matrix T, and where h_1 is an $(n-k) \times (n-k)$ matrix satisfying $h_1^* = h_1$ and h_2 is a $k \times k$ matrix satisfying $h_2^* = h_2$. We have $(i(e_{11} - e_{22}), i(e_{11} - e_{22})) < 0$ whether $k = 0$, $k = 1$, or $k \geqslant 2$, and thus $\operatorname{tr}(\lambda\lambda')$ is, once again, a positive multiple of the natural bilinear form. Using the above partitioning of the matrices of $HA^{(k)}_{n-1}$ we see easily that the signature of $HA^{(k)}_{n-1}$ is $2(n-k)k$; thus all the algebras $HA^{(k)}_{n-1}$, $0 \leqslant k \leqslant n/2$, are non-isomorphic.

We have noted just above that for $n > 2$ no conjugation $\lambda \to T\bar{\lambda}T^{-1}$ is equivalent to any conjugation $\lambda \to -T\bar{\lambda}'T^{-1}$; thus neither of the groups RA_{n-1} or QA_{n-1} is isomorphic to any of the groups $HA^{(k)}_{n-1}$. In the degenerate case $n = 2$ however, we obtain only the groups RA_1 and QA_1.

This completes our analysis of the conjugations in A_{n-1}. We now turn to an analysis of the conjugations in the Lie algebras of types B_m, C_m, and D_m, which, because of their similarity, we give in parallel.

B. Analysis of the conjugations in the Lie algebras of types B_m, C_m, and D_m

We recall from section II.4 that B_m and D_m are the Lie algebras of all $p \times p$ complex matrices λ of trace zero such that $\langle\langle \lambda x, y\rangle\rangle = -\langle\langle x, \lambda y\rangle\rangle$, where

$$\langle\langle x, y\rangle\rangle = x_1 y_p + \cdots + x_p y_1 . \tag{2}$$

Here and below, $p = 2m + 1$ if we deal with B_m, and $p = 2m$ if we deal with D_m. We may describe C_m in just the same way, by putting $p = 2m$, and writing

$$\langle\langle x, y\rangle\rangle = x_1 y_p + \cdots + x_m y_{m+1} - x_{m+1} y_m - \cdots - x_p y_1 \tag{3}$$

instead of (2).

As we saw in the course of proving Lemma 9, the natural bilinear form (λ, μ) is in each case a positive multiple of tr $(\lambda\mu)$.

The Cartan subalgebra $\hat{M}$ of these Lie algebras $\hat{\Lambda}$ is in all cases the collection of all diagonal matrices of trace zero. In all cases it follows readily by Definition II.5.3 and Lemma I.5.3 that an inner automorphism has the form $\lambda \to T\lambda T^{-1}$, where $\langle\langle Tx, Ty\rangle\rangle = \langle\langle x, y\rangle\rangle$, and where $\det(T) = +1$. By Lemma 12, all automorphisms of B_m and C_m are inner. If $m > 4$, then, again by Lemma 12, the number of cosets of the group A_0 of inner automorphisms of D_m in the group A of all automorphisms is at most 2; if $m = 4$, A/A_0 has at most six elements. The Lie algebra D_m admits the automorphism $\lambda \to R\lambda R^{-1}$, where R is the transformation satisfying $\langle\langle Rx, Ry\rangle\rangle = \langle\langle x, y\rangle\rangle$ and $\det(R) = -1$ defined by

$$R[x_1, \ldots, x_{2m}] = [x_1, \ldots, x_{m-1}, x_{m+1}, x_m, x_{m+2}, \ldots, x_{2m}]. \tag{4}$$

Now, if $S\lambda S^{-1} = \lambda$ for all $\lambda \in \hat{\Lambda}$ then the eigenspaces of S are invariant under all the transformations $\lambda \in \hat{\Lambda}$, and it follows readily that S is a constant multiple of the identity. If in addition $\langle\langle Sx, Sy\rangle\rangle = \langle\langle x, y\rangle\rangle$, we must have $S = \pm I$. Using this, we see that if the automorphism $\lambda \to R\lambda R^{-1}$ of D_m were inner, there would exist T satisfying $\langle\langle Tx, Ty\rangle\rangle = \langle\langle x, y\rangle\rangle$ and $\det(T) = 1$, such that $RT^{-1} = \pm I$. Taking determinants. we would get $\det(R) = 1$, a contradiction; which proves that $\lambda \to R\lambda R^{-1}$ is not inner. We conclude that for D_m, $m > 4$, A/A_0 has 2 components: every automorphism is of the form $\lambda \to T\lambda T^{-1}$ with $\langle\langle Tx, Ty\rangle\rangle = \langle\langle x, y\rangle\rangle$; if $\det(T) = +1$ we have an inner automorphism, if $\det(T) = -1$ we have an outer automorphism.

If λ' denotes the transpose of the matrix λ, we have $\lambda' = -J\lambda J^{-1}$, where $J = J_+$ is the matrix of the transformation $[x_1, \ldots, x_p] \to [x_p, \ldots, x_1]$ if we deal with B_m or D_m, and $J = J_-$ is the matrix of the transformation $[x_1, \ldots, x_{2m}] \to [x_{2m}, \ldots, x_{m+1}, -x_m, \ldots, -x_1]$ if we deal with C_m. The positive conjugation C of Lemma 9 is $\lambda \to J\bar{\lambda}J^{-1}$.

It follows from the above that, except for the type D_4, every conjugation is of the form $\lambda \to TJ\bar{\lambda}J^{-1}T^{-1}$, where $\langle\langle Tx, Ty\rangle\rangle = \langle\langle x, y\rangle\rangle$. We shall now show that, even for the type D_4, every conjugation is equivalent to a conjugation of the form $\lambda \to TJ\bar{\lambda}J^{-1}T^{-1}$, where $\langle\langle Tx, Ty\rangle\rangle = \langle\langle x, y\rangle\rangle$. We argue as follows. As we saw in section 4, the roots of $\hat{M}$ in $\hat{\Lambda}$ (of type D_4) are $\pm(\omega_i - \omega_j)$ and $\pm(\omega_i + \omega_j)$, $i \neq j$, $1 \leqslant i, j \leqslant 4$. The roots $\omega_1 - \omega_2, \omega_2 - \omega_3, \omega_3 - \omega_4, \omega_3 + \omega_4$, which we shall call $\alpha_1, \alpha_4, \alpha_2, \alpha_3$ in the remainder of our discussion of the Lie algebra of type D_4, form a simple set Π. Any permutation π of the roots $\alpha_1, \alpha_2, \alpha_3$ is a symmetry of the Dynkin diagram. Thus, by Theorem II.5.4 and Corollary II.3.22, if we put $\pi(4) = 4$ for notational convenience, the mapping $\lambda_{\pm\alpha_i} \to \lambda_{\pm\alpha_{\pi(i)}}$ extends in a unique way to an automorphism $[\pi]$ of $\hat{\Lambda}$. This gives a group P of six automorphisms of $\hat{\Lambda}$. By Theorem II.5.1 and the remark which follows its statement, and by Theorem II.5.4, every automorphism ϕ of $\hat{\Lambda}$ has the form $\phi = \psi[\pi]$, where ψ is inner, and $[\pi] \in P$. Define the conjugate $\bar{\phi}$ of an arbitrary automorphism ϕ by $\phi(\lambda)$ $= \overline{\phi(\bar{\lambda})}$, $\lambda \in \hat{\Lambda}$. Then it is easily verified that $\bar{\phi}$ is an inner automorphism if ϕ is an inner automorphism. The description of the root matrices corresponding to the various non-zero roots of $\hat{M}$ in $\hat{\Lambda}$ given in Section II.4 makes it plain that the matrices λ_α, $0 \neq \alpha \in \{\alpha\}$, may be taken to be real; thus $[\overline{\pi}] = [\pi]$, $[\pi] \in P$. Every conjugation C_1 in $\hat{\Lambda}$ is of the form $C_1\lambda \to \phi(\bar{\lambda})$, where $\phi \in A$ and ϕ satisfies $\bar{\phi}\phi = I$. Using the fact that $[\pi_1][\pi_2] = [\pi_1\pi_2]$ for $[\pi_1], [\pi_2] \in P$, we see that the conjugation C_1 is either $\lambda \to \psi(\bar{\lambda})$ where ψ is inner, or is equivalent to a conjugation $\lambda \to (\psi[(23)])(\bar{\lambda})$ where ψ is inner, or is equivalent to a conjugation $\lambda \to (\psi[(123)])(\bar{\lambda})$ where ψ is inner. Now, if $[(123)]$ is not inner, neither is $[(132)]$, since $[(132)]^2 = [(123)]$. But then, since $\overline{\psi[(123)]}\,\psi[(123)] = \psi\psi_1[(132)]$ where $\psi, \psi_1 \in A_0$, $\psi[(123)]\,\overline{\psi[(123)]} \neq I$. Therefore every conjugation is either of the form $\lambda \to \psi(\bar{\lambda})$, where ψ is inner, or is equivalent to a conjugation $\lambda \to (\psi[(23)])(\bar{\lambda})$, where ψ is inner. Define the image of a root α under an automorphism ϕ mapping M into itself by $(\phi\alpha)(\mu) = \alpha(\phi^{-1}\mu)$, $\mu \in \hat{M}$; or, equivalently, by $\phi(\mu_\alpha) = \mu_{\varphi(\alpha)}$. Then the automorphism $[(23)]$ maps $\omega_1 - \omega_2, \omega_2 - \omega_3, \omega_3 - \omega_4, \omega_3 + \omega_4$ into $\omega_1 - \omega_2, \omega_2 - \omega_3, \omega_3 + \omega_4, \omega_3 - \omega_4$. According to the definition of the linear functions ω_i given in section 4, we have $\omega_i(e_{jj} - e_{9-j,9-j}) = \delta_{ij}$, $j = 1, \ldots, 4$. Thus [(23)] has the same effect on each vector of M as the automorphism $\lambda \to R\lambda R^{-1}$, where R is the matrix of (4). Our assertion that every conjugation of $\hat{\Lambda}$ is equivalent to a conjugation of the form $\lambda \to TJ\bar{\lambda}J^{-1}T^{-1}$ where $\langle\langle Tx, Ty\rangle\rangle = \langle\langle x, y\rangle\rangle$ is therefore proved.

It now follows, using Corollary 16, that in all cases every conjugation of Λ is equivalent to a conjugation of the form $\lambda \to TJ\bar{\lambda}J^{-1}T^{-1}$, where $\langle\langle Tx, Ty\rangle\rangle = \langle\langle x, y\rangle\rangle$, $T^2 = \pm I$, and $JT'J = \pm T$. We analyze the two cases $T^2 = I$ and $T^2 = -I$ separately.

(+) If $T^2 = I$, the eigenvalues of T are ± 1, and the space V of all vectors decomposes as the direct sum $V = V_{+1} \oplus V_{-1}$ of the two corresponding eigenspaces of V. We may assume, replacing T by $-T$ if necessary, that V_{+1} has the larger dimension. From $\langle\langle Tx, Ty\rangle\rangle = \langle\langle x, y\rangle\rangle$ we deduce that $\langle\langle V_{+1}, V_{-1}\rangle\rangle = 0$. An evident induction now shows that in case we deal either with B_m or with D_m, so that $\langle\langle x, y\rangle\rangle = \langle\langle y, x\rangle\rangle$, we may construct bases $x_1, \ldots, x_q$ for V_{+1} and $y_1, \ldots, y_r$ for V_{-1} such that $\langle\langle x_i, x_j\rangle\rangle = \delta_{ij} = \langle\langle y_i, y_j\rangle\rangle$, $\langle\langle x_i, y_j\rangle\rangle = 0$; if we deal with C_m, so that $\langle\langle x, y\rangle\rangle = -\langle\langle y, x\rangle\rangle$, we find similarly that the dimension of V_{+1} must be even, and that there exist bases $x_1, x_1', x_2, x_2', \ldots, x_q, x_q'$ for V_{+1}, and $y_1, y_1', \ldots, y_r, y_r'$ for V_{-1}, such that $\langle\langle x_i, x_j\rangle\rangle = \langle\langle x_i', x_j'\rangle\rangle = \langle\langle y_i, y_j\rangle\rangle = \langle\langle y_i', y_j'\rangle\rangle = 0$, $\langle\langle x_i, x_j'\rangle\rangle = \delta_{ij} = \langle\langle y_i, y_j'\rangle\rangle$. It is therefore clear, using Corollary 18, that if S is any other transformation such that $S^2 = I$, $\langle\langle Sx, Sy\rangle\rangle = \langle\langle x, y\rangle\rangle$, and $J\bar{S}J = \pm S$, and if the dimension of the eigenspace of S belonging to the eigenvalue $+1$ is the same as the dimension of V_{+1}, then the conjugations $\lambda \to TJ\bar{\lambda}J^{-1}T^{-1}$ and $\lambda \to SJ\bar{\lambda}J^{-1}S^{-1}$ are equivalent. We may take S to be the matrix R_k of the transformation

$$[x_1, \ldots, x_p] \to [x_p, x_{p-1}, \ldots, x_{p+1-k}, x_{k+1}, \ldots, x_{p-k}, x_k, \ldots, x_1]$$

where $k \leqslant m$ if we deal with B_m or D_m, and S to be the matrix Δ_k of the transformation

$$[x_1, \ldots, x_p] \to [x_1, \ldots, x_k, -x_{k+1}, \ldots, -x_{p-k}, x_{p+1-k}, \ldots, x_p]$$

where $k \leqslant m/2$ if we deal with C_m.

(−) If $T^2 = -I$, the eigenvalues of T are $\pm i$, and the space V of vectors decomposes as the direct sum $V = V_i \oplus V_{-i}$. We have $\langle\langle V_i, V_i\rangle\rangle = 0 = \langle\langle V_{-i}, V_{-i}\rangle\rangle$. By induction, we may construct bases $x_1, \ldots, x_m$ for V_i and $y_1, \ldots, y_m$ for V_{-i}, such that $\langle\langle x_j, y_k\rangle\rangle = \delta_{jk}$. In particular, the case $T^2 = -I$ can never occur if p is odd, i.e., if we deal with B_m. We find as above that our conjugation is equivalent to the conjugation

$\lambda \to i\Delta'_m J\bar{\lambda}J^{-1}\,(i\Delta'_m)^{-1}$ in both cases, where Δ'_m is the matrix of the transformation

$$[x_1, \ldots, x_{2m}] \to [x_1, \ldots, x_m, -x_{m+1}, \ldots, -x_{2m}].$$

All in all, we find the following real Lie algebras:

$LB_m^{(k)}$, the set of all $\lambda \in B_m$ such that $R_k J\bar{\lambda}J^{-1}R_k^{-1} = \lambda;\ 0 \leqslant k \leqslant m;$

$HC_m^{(k)}$, the set of all $\lambda \in C_m$ such that $\Delta_k J\bar{\lambda}J^{-1}\Delta_k^{-1} = \lambda;\ 0 \leqslant k \leqslant \dfrac{m}{2};$

RC_m, the set of all $\lambda \in C_m$ such that $\Delta'_m J\bar{\lambda}J^{-1}\Delta_m'^{-1} = \lambda;$

$LD_m^{(k)}$, the set of all $\lambda \in D_m$ such that $R_k J\bar{\lambda}J^{-1}R_k^{-1} = \lambda;\ 0 \leqslant k \leqslant m;$

QD_m, the set of all $\lambda \in D_m$ such that $\Delta'_m J\bar{\lambda}J^{-1}\,(\Delta'_m)^{-1}$.

We shall now calculate the signatures of those various Lie algebras, and show that no two of them are isomorphic. We begin with the algebras $LB_m^{(k)}$ and $LD_m^{(k)}$. Suppose that we renumber the coordinates so that $x_1, \ldots, x_k$ and $x_{p-k+1}, \ldots, x_p$ appear as $y_1, \ldots, y_{2k}$, and the remaining x's as $y_{2k+1}, \ldots, y_p$, and partition the re-indexed matrices $\lambda \in \hat{\Lambda}$ accordingly. Then, letting j denote a matrix of the appropriate dimension having the same structure as J_+, the re-indexed matrix J_+ will be partitioned as

$$J = \begin{pmatrix} j & \\ & j \end{pmatrix},$$

and the matrix $R_k J$ will appear as

$$R_k J = \begin{pmatrix} I & \\ & j \end{pmatrix}.$$

Expressing the condition $\langle\langle \lambda x, y\rangle\rangle = -\langle\langle x, \lambda y\rangle\rangle$ as $J\lambda = -\lambda' J$, we find that the matrices λ in the real Lie algebras $LB_m^{(k)}$, $HC_m^{(k)}$, and $LD_m^{(k)}$ appear as

$$\begin{pmatrix} r & b \\ -jb'j & ih \end{pmatrix},$$

where r is real and satisfies $-r' = jrj^{-1}$ but is otherwise arbitrary, b satisfies $b' = bj^{-1}$ but is otherwise arbitrary, and $h = \bar{h}'$. The expression (λ, λ) is then a positive multiple of $\operatorname{tr}(r^2 - 2\overline{b'b} - h^2)$. The reader will then verify that the term $\operatorname{tr}(r^2)$ contributes as many positive squares as there are pairs i, j with $2k - j \geqslant i \geqslant j$, which is to say k^2, while the term

$-\mathrm{tr}\,(\overline{b'b})$ contributes a number of positive squares equal to $k(p-2k)$; the signature of $LB_m^{(k)}$ and $LD_m^{(k)}$ is thus $k(p-k)$.

To calculate the signature of the Lie algebras $HC_m^{(k)}$, we partition the matrices $\lambda \in HC_m^{(k)}$ in a manner corresponding to the threefold partition of $x_1, \ldots, x_{2m}$ into $x_1, \ldots, x_k; x_{k+1}, \ldots, x_{p-m}; x_{p-m+1}, \ldots, x_p$, and use the equations $\lambda = -\Delta_k \bar{\lambda}' \Delta_k$ and $\lambda' = -J\lambda J^{-1}$. We find that λ has the form

$$\lambda = \begin{pmatrix} \xi_1 & \zeta_1 & \xi_4 \\ \zeta_1^* & \xi_2 & \zeta_2 \\ -\xi_4^* & \zeta_2^* & \xi_3 \end{pmatrix},$$

where as usual we write ζ^* for the conjugate transpose of the matrix ζ. In the matrix displayed above, $\xi_1 = -\xi_1^*$, $\xi_2 = -\xi_2^*$, and $\xi_3 = -\xi_3^*$, while $(\zeta_1)_{ij} = \pm(\zeta_2)_{p+1-i\; p+1-j}$, but ζ_1 is otherwise arbitrary. But then $\mathrm{tr}\,(\lambda^2) = -\mathrm{tr}\,(\xi_1\xi_1^* + \xi_2\xi_2^* + \xi_3\xi_3^* + \xi_4\xi_4^*) + 2\,\mathrm{tr}\,(\zeta_1\zeta_1^*)$, making it clear that the signature of $HC_m^{(k)}$ is $2k\,(2m-2k)$.

To calculate the signature of RC_m, we partition each $\lambda \in RC_m$ as

$$\lambda = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

and note that $\bar{\lambda} = -\Delta_m' \lambda' \Delta_m'$, and that $\lambda' = -J\lambda J^{-1}$. Thus $a = -a^*$, $d = -d^*$, and $b = c^*$, while $(b)_{m+1-i,m+1-j} = (b)_{ij}$ and b is otherwise arbitrary. The number of positive squares in $\mathrm{tr}\,(\lambda^2) = \mathrm{tr}\,(a^2 + d^2 + 2bb^*)$ is then $m(m+1)$. In exactly the same way, we find that the signature of QD_m is $m(m-1)$. Examining the signature, we see that the only possible isomorphisms among the above real Lie algebras arise when we deal with $LD_m^{(k)}$ and QD_m, and when $k^2 - 2km + m(m-1) = 0$, i.e., m is a perfect square, and $k = m - \sqrt{m}$. Let us first consider the cases $m > 4$. In these cases, we have seen above that every automorphism of D_m has the form $\lambda \to T\lambda T^{-1}$, where T is such that $\langle\langle Tx, Ty\rangle\rangle = \langle\langle x, y\rangle\rangle$. Thus, if $\lambda \to R_k J\bar{\lambda}J^{-1}R_k^{-1}$ and $\lambda \to \Delta_m' J\bar{\lambda}J^{-1}(\Delta_m')^{-1}$ were equivalent conjugations, there would have to exist such a matrix T satisfying $TR_kJ\bar{T}^{-1} = \pm\,\Delta_m' J$. But this would imply that $\bar{T}R_kJT^{-1}TR_kJ\bar{T}^{-1} = \Delta_m' J \Delta_m' J$, i.e., $I = -I$, which is impossible.

We may obtain the same conclusion for $m = 4$, at the cost of a little extra labor, as follows. We have seen above that there exists a group of six automorphisms $[\pi]$ of D_4, isomorphic to the permutation group of

3 objects, such that every automorphism of D_4 has the form $\psi[\pi]$, where ψ is inner. Moreover, as we have seen, $[(23)]$ is not inner, so that $[(12)]$ and $[(13)]$ are not inner either. Products of the form $[(23)]\,[(12)]^{-1}$ will be inner if and only if the automorphisms $[(123)]$ and $[(132)]$ are inner. All the conjugations which we have considered above have the form $\lambda \to \psi[(23)]\,\bar{\lambda}$, where ψ is inner. If two such conjugations $C_1 : \lambda \to \psi_1[(23)]\,\bar{\lambda}$ and $C_2 : \lambda \to \psi_2[(23)]\,\bar{\lambda}$ are equivalent, there would exist an inner automorphism ϕ and a permutation π such that $\psi_1[(23)] = \phi[\pi]\psi_2\,[(23)]\bar{\phi}^{-1}\,[\pi]$. But the right hand side has the form $\phi_1[\pi(23)\pi^{-1}]$, where ϕ_1 is inner. Thus $[(23)]\,[(\pi(2)\,\pi(3))]^{-1}$ would be inner. If $[(123)]$ is not inner, this means that the permutation π must map the set $\{2, 3\}$ into itself, so that either $\pi = e$ or $\pi = (23)$. If (123) is inner, however, it follows that *every* automorphism is of the form η or $\eta[(23)]$, where η is inner. Therefore, in both possible cases, if the two conjugations C_1 and C_2 are equivalent, there exists a transformation T such that $\langle\langle Tx, Ty\rangle\rangle = \langle\langle x, y\rangle\rangle$: and such that $C_1\lambda = TC_2T^{-1}\lambda$. With this conclusion in hand, we may argue in the case $m = 4$ exactly as in the cases $m > 4$, to conclude that $LD_4^{(2)}$ and QD_4 are not isomorphic.

Our analysis of the "classical" types of the real Lie algebras is now complete. The following theorem summarizes the results which we have obtained.

21. Theorem. The real simple Lie algebras whose complexifications are simple and of one of the "classical" forms A_m, $m \geqslant 1$; B_m, $m \geqslant 2$; C_m, $m \geqslant 3$; or D_m, $m \geqslant 4$, are as follows:

(a) If $m > 1$, the algebras $HA_m^{(k)}$, $0 \leqslant k \leqslant \dfrac{m+1}{2}$, of signature $2(m + 1 - k)\,k$; the algebra RA_m of signature $\frac{1}{2}m\,(m + 3)$, and, if $m > 1$ is odd, the algebra QA_m of signature $\frac{1}{2}(m + 2)\,(m - 1)$. If $m = 1$, the algebra RA_m of signature 2 and the algebra QA_m of signature 0.

(b) The algebras $LB_m^{(k)}$, $0 \leqslant k \leqslant m$, of signature $k(2m + 1 - k)$.

(c) The algebras $HC_m^{(k)}$, $0 \leqslant k \leqslant \dfrac{m}{2}$, of signature $2k\,(2m - 2k)$; the algebra RC_m of signature $m(m + 1)$.

(d) The algebras $LD_m^{(k)}$, $0 \leqslant k \leqslant m$, of signature $k(2m - k)$; the algebra QD_m, of signature $m(m - 1)$.

None of these real Lie algebras are isomorphic. A model for each of them is given in the preceding paragraphs.

To prepare for the study of the conjugations in the five exceptional Lie algebras, we first develop some general theoretical material, which, like

Lemma 14, extends certain important theorems of linear algebra to the group of automorphisms of an arbitrary semi-simple complex Lie algebra.

22. Lemma. "Jordan Canonical Form". Let ϕ be an automorphism in the complex Lie algebra $\tilde{\Lambda}$. Then ϕ may be factored as $\phi = \psi \exp(\tau)$, where

(i) ψ and τ commute,
(ii) ψ is an automorphism of $\hat{\Lambda}$,
(iii) there exists a complete basis for $\hat{\Lambda}$ consisting of eigenvectors of the automorphism ψ,
(iv) $\tau[\lambda, \mu] = [\tau\lambda, \mu] + [\lambda, \tau\mu]$, $\lambda, \mu \in \hat{\Lambda}$, and
(v) τ is a nilpotent transformation in the vector space $\hat{\Lambda}$.

Proof: Let $\{\beta\}$ be the set of eigenvalues of ϕ in $\hat{\Lambda}$, and let $\Lambda_{(\beta)}$ be the corresponding generalized eigenspaces. If $\lambda \in \Lambda_{(\beta)}$ and $\lambda' \in \Lambda_{(\beta')}$, we have

$$\begin{aligned}(\phi - \beta\beta')[\lambda, \lambda'] &= [\phi\lambda, \phi\lambda'] - [\beta\lambda, \beta'\lambda'] \\ &= [\phi\lambda, \phi\lambda'] - [\phi\lambda, \beta'\lambda'] + [\phi\lambda, \beta'\lambda'] - [\beta\lambda, \beta'\lambda'] \\ &= [\phi\lambda, (\phi - \beta')\lambda'] + [(\phi - \beta)\lambda, \beta'\lambda'].\end{aligned}$$

Starting with this identity and proceeding inductively we find that the quantity $(\phi - \beta\beta')^p[\lambda, \lambda']$ may be written as a sum of terms of the form $[\phi^i(\phi - \beta)^j\lambda, (\beta')^k(\phi - \beta')^l\lambda'], j + l \geqslant p$. It follows that $(\phi - \beta\beta')^p[\lambda, \lambda'] = 0$ for sufficiently large p, i.e., that $[\Lambda_{(\beta)}, \Lambda_{(\beta')}] \subseteq \Lambda_{(\beta\beta')}$.

Thus, if we define a linear transformation ψ on $\hat{\Lambda}$ by putting $\psi\lambda = \beta\lambda$ if $\lambda \in \Lambda_{(\beta)}$, we have $[\psi\lambda, \psi\lambda'] = \psi[\lambda, \lambda']$ for $\lambda \in \Lambda_{(\beta)}$ and $\lambda' \in \Lambda_{(\beta')}$, and therefore, since $\hat{\Lambda} = \sum_{\beta \in \{\beta\}} \oplus \Lambda_{(\beta)}$, $[\psi\lambda, \psi\lambda'] = \psi[\lambda, \lambda']$ for all $\lambda, \lambda' \in \hat{\Lambda}$, so that ψ is an automorphism of $\hat{\Lambda}$. It follows that $\eta = \phi\psi^{-1}$ is an automorphism of $\hat{\Lambda}$; plainly, $(\eta - I)^p = 0$ for sufficiently large p. Since ϕ plainly commutes with ψ, so does η.

Define the linear transformation τ in $\hat{\Lambda}$ by

$$\tau = \log \eta = (I - \eta) - \tfrac{1}{2}(I - \eta)^2 + \tfrac{1}{3}(I - \eta)^3 + \cdots;$$

(note that the series actually terminates). Then it is plain that τ commutes with ψ, and that $\tau^p = 0$ for sufficiently large p. Then $\exp(n\tau) = \eta^n$ for each integer n by Lemma 0.35. Therefore, if γ is any linear functional on $\hat{\Lambda}$ and $\lambda, \mu \in \hat{\Lambda}$, the complex valued function

$$\gamma(\exp(z\tau)[\lambda, \mu] - [\exp(z\tau)\lambda, \exp(z\tau)\mu])$$

vanishes whenever z is an integer. On the other hand, since $\tau^p = 0$, we see from the exponential series that this function is actually a polynomial in z. Hence, it vanishes identically, i.e. we have identically

$$\exp(z\tau)\,[\lambda, \mu] = [\exp(z\tau)\,\lambda, \exp(z\tau)\,\mu].$$

Differentiating with respect to z and putting $z = 0$, we obtain our lemma. Q.E.D.

23. Corollary. Let ϕ be an automorphism in the complex semi-simple Lie algebra $\hat{\Lambda}$. Then ϕ may be factored as $\phi = \psi \exp(\mathrm{Ad}(\lambda_0))$, where $\lambda_0 \in \hat{\Lambda}$, $\psi\lambda_0 = \lambda_0$, $\mathrm{Ad}(\lambda_0)$ is a nilpotent operator, and there exists a complete basis for $\hat{\Lambda}$ consisting of eigenvectors of the automorphism ψ.

Proof. By Lemma 13, we may write the mapping τ of the preceding lemma as $\tau = \mathrm{Ad}(\lambda_0)$, where $\lambda_0 \in \hat{\Lambda}$. We have

$$0 = \psi\,\mathrm{Ad}(\lambda_0) - \mathrm{Ad}(\lambda_0)\,\psi = \psi\,\mathrm{Ad}(\lambda_0 - \psi^{-1}\lambda_0),$$

and thus $\psi\lambda_0 = \lambda_0$ by Theorem 2.3. Q.E.D.

24. Corollary. Let ϕ be an automorphism in a complex Lie algebra $\hat{\Lambda}$, and let $\{\beta\}$ be the set of eigenvalues of ϕ, and let $\Lambda_{(\beta)}$ be the corresponding generalized eigenspaces. Then $[\Lambda_{(\beta)}, \Lambda_{(\beta')}] \subseteq \Lambda_{(\beta\beta')}$ if $\beta, \beta' \in \{\beta\}$.

Proof: This follows as in the first paragraph of the proof of the preceding lemma. Q.E.D.

Next we single out a special class of automorphisms "with simple characteristic roots".

25. Definition. An automorphism ϕ of a complex Lie algebra $\hat{\Lambda}$ is said to *regular* if there exists a neighborhood U in the group of all automorphisms of $\hat{\Lambda}$ such that $\phi \in U$, and such that no automorphism $\psi \in U$ has (when regarded as a linear transformation in $\hat{\Lambda}$) more eigenvalues than ϕ.

Note that when in the preceding definition we speak of neighborhoods, we mean to regard the group of automorphisms of $\hat{\Lambda}$ as a closed subgroup of the Lie group of all non-singular linear transformations of $\hat{\Lambda}$.

26. Lemma. Let ϕ be a regular automorphism of the complex semi-simple algebra $\hat{\Lambda}$. Then there exists a complete basis for $\hat{\Lambda}$ consisting of eigenvectors of ϕ.

Proof. Let $\{\beta\}$ be the set of eigenvalues of ϕ, and let $\Lambda_{(\beta)}$ be the corresponding generalized eigenspaces. By Corollary 24, $\Lambda_{(1)}$ is an algebra of $\hat{\Lambda}$. Our first aim is to prove that $\Lambda_{(1)}$, regarded as a subalgebra in its own right, is nilpotent. We use an argument closely related to that used to prove Theorem II.2.8, as follows.

By Corollary 23, we may write $\phi = \psi \exp(\mathrm{Ad}(\lambda_0))$, where $\psi|\Lambda_{(\beta)} = \beta I$ for each $\beta \in \{\beta\}$, where $\mathrm{Ad}(\lambda_0)$ is nilpotent, and where $\phi(\lambda_0) = \lambda_0$, so that $\lambda_0 \in \Lambda_{(1)}$. Let $\lambda \in \Lambda_{(1)}$. Then, if ζ is close to 1 and z is close to zero but non-zero, the automorphism $\phi_1 = \psi \exp(\mathrm{Ad}(\zeta(\lambda_0 + z\lambda)))$ is close to ϕ. By Corollary 24, this automorphism maps each space $\Lambda_{(\beta)}$ into itself. By Lemma 0.27, $\phi_1|\Lambda_{(\beta)}$ has an eigenvalue near β. Thus, since ϕ is regular, $\phi_1|\Lambda_{(\beta)}$ can have at most one eigenvalue for each $\beta \in \{\beta\}$. Since $\psi|\Lambda_{(1)} = I$, and since $\exp(\mathrm{Ad}(\zeta(\lambda_0 + z\lambda)))$ always admits $\lambda_0 + z\lambda \in \Lambda_{(1)}$ as an eigenvector belonging to the eigenvalue 1, the only eigenvalue of $\exp(\mathrm{Ad}(\zeta(\lambda_0 + z\lambda)))|\Lambda_{(1)}$ is 1. It follows by Lemma 0.34 that all the eigenvalues of $\mathrm{Ad}(\zeta(\lambda_0 + z\lambda))|\Lambda_{(1)}$ are of the form $2\pi i n$, where n is an integer. Since ζ can be any complex number close to 1, it follows that the only eigenvalue of $\mathrm{Ad}(\lambda_0 + z\lambda)|\Lambda_{(1)}$ is zero. Thus $(\mathrm{Ad}(\lambda_0 + z\lambda)|\Lambda_{(1)})^p = 0$ for some finite p and for all sufficiently small non-zero z. But then $(\mathrm{Ad}(\lambda_0 + z\lambda)|\Lambda_{(1)})^p = 0$ for all z; multiplying by z^{-p} and letting $|z| \to \infty$, we find that $(\mathrm{Ad}(\lambda)|\Lambda_{(1)})^p = 0$ for all $\lambda \in \Lambda_{(1)}$. Thus, by Theorem II.1.11, the Lie algebra $\Lambda_{(1)}$ is nilpotent.

Let $\lambda \in \Lambda_{(1)}$, and $\mu \in \Lambda_{(\beta)}$ with $\beta \neq 1$. By Corollary 24, $\mathrm{Ad}(\lambda)\,\mathrm{Ad}(\mu)$ maps $\Lambda_{(\beta')}$ into $\Lambda_{(\beta\beta')}$. If we form a basis for $\hat{\Lambda}$ out of bases for each of the separate spaces $\Lambda_{(\beta')}$, it follows that the matrix $\mathrm{Ad}(\lambda)\,\mathrm{Ad}(\mu)$ in this basis has no non-zero diagonal entries; hence $\mathrm{tr}(\mathrm{Ad}(\lambda)\,\mathrm{Ad}(\mu)) = 0$ for $\lambda \in \Lambda_{(1)}$ and $\mu \in \Lambda_{(\beta)}$, $\beta \neq 1$. It follows by Theorem II.2.3 that if $\lambda \in \Lambda_{(1)}$ and $\lambda \neq 0$, there exists a $\lambda' \in \Lambda_{(1)}$ such that $\mathrm{tr}(\mathrm{Ad}(\lambda)\,\mathrm{Ad}(\lambda')) \neq 0$.

Let $\{\gamma\}$ be the set of roots of $\Lambda_{(1)}$ in $\hat{\Lambda}$, and $\hat{\Lambda} = \sum_{\gamma \in \{\gamma\}} \oplus \Lambda^{(\gamma)}$ the corresponding decomposition of $\hat{\Lambda}$ into root spaces. Let $\lambda, \lambda' \in \Lambda_{(1)}$. It follows from Corollary II.1.13 that $\Lambda^{(\gamma)}$ is a generalized eigenspace of $\mathrm{Ad}(\lambda)\,\mathrm{Ad}(\lambda')$ belonging to the eigenvalue $\gamma(\lambda)\gamma(\lambda')$; thus $\mathrm{tr}(\mathrm{Ad}(\lambda)\,\mathrm{Ad}(\lambda')) = \sum_{\gamma \in \{\gamma\}} \gamma(\lambda)\gamma(\lambda')$ Therefore, if $\lambda \neq 0$, there must exist some $\gamma \in \{\gamma\}$ such that $\gamma(\lambda) \neq 0$. But, since $\mathrm{Ad}(\lambda_0)$ is nilpotent, $\gamma(\lambda_0) = 0$ for all γ. Thus $\lambda_0 = 0$, and it follows that $\phi = \psi$, so that there exists a complete basis for $\hat{\Lambda}$ consisting of eigenvectors of ϕ. Q.E.D.

27. Definition. Let T_1 and T_2 be linear transformations in a vector space V. Then T_1 and T_2 *commute totally* if every linear transformation T in V which commutes with T_1 also commutes with T_2, and vice-versa.

Note that the relation thus defined is symmetric, reflexive and transitive, and is therefore an equivalence relation.

28. Corollary. Let ϕ be a regular automorphism of the complex semisimple algebra $\hat{\Lambda}$. Then every automorphism ψ which is sufficiently close to ϕ and which commutes with ϕ commutes totally with ϕ.

Proof. Let $\{\beta\}$ be the eigenvalues of ϕ, and let $\Lambda_{(\beta)}$ be the corresponding eigenspaces; we have $\phi|\Lambda_{(\beta)} = \beta I$ amd $\hat{\Lambda} = \sum_{\beta\{\in\beta\}} \oplus \Lambda_{(\beta)}$ by Lemma 26. Plainly, $\psi\Lambda_{(\beta)} \subseteq \Lambda_{(\beta)}$. If ψ is sufficiently close to ϕ, then, since ϕ is regular, $\psi|\Lambda_{(\beta)}$ can have at most one eigenvalue $\gamma_{(\beta)}$, and this eigenvalue must lie close to β. It follows at once from Definition 25 that ψ is regular, and thus, by Lemma 26, that $\psi|\Lambda_{(\beta)} = \gamma_{(\beta)}I$. It is then quite plain that every linear transformation in $\hat{\Lambda}$ which commutes with ψ also commutes with ϕ, and vice-versa, so that ϕ and ψ commute totally. Q.E.D.

29. Lemma. Let $\hat{\Lambda}$ be a complex Lie algebra, A its group of automorphisms, and A' a connected component of A. Then, if A'' is the set of all regular elements of A', A'' is connected.

Proof. Since A is a closed subgroup of the group of all non-singular linear transformations in $\hat{\Lambda}$, it follows by Theorem I.6.12 and Lemma I.4.11 that there exists a set $\mathfrak{L}$ of linear transformations closed under addition and under multiplication by real numbers, a neighborhood U of the identity in A, and a neighborhood V of the zero element in $\mathfrak{L}$, such that $L \to \exp(L)$ is a homeomorphism of V on U. If $L \in \mathfrak{L}$, then $\exp(tL)$ is an automorphism of $\hat{\Lambda}$ for all sufficiently small real t, and therefore for all complex t; thus, if z is complex and sufficiently small, we have $\exp(zL) = \exp(L')$, where $L' \in \mathfrak{L}$ and is small. Thus, since we see on applying Lemma I.4.11 once more that the map $L_0 \to \exp(L_0)$ is a homeomorphism of a neighborhood of zero in the space of all matrices onto a neighborhood of I in the group of non-singular matrices, we must have $zL = L'$. This shows that $zL \in \mathfrak{L}$ if $L \in \mathfrak{L}$ and z is complex.

Let $\phi \in A'$, and let ϕ be regular and have n eigenvalues. We shall show next that no $\psi \in A'$ has more than n eigenvalues, and that if $\psi \in A'$ and has exactly n eigenvalues, then ψ is regular. To do this, define $m(\psi)$ for each $\psi \in \Lambda$ as the smallest integer m for which there exists a neighborhood U of ψ in A such that any $\eta \in U$ has at most m eigenvalues. Plainly, each of the sets $A_k = \{\psi \in A|m(\psi) \leqslant k\}$ is open. We assert that each of the sets A_k is closed. Indeed, suppose the contrary, i.e., that $n(\psi) > k$, but that ψ is the limit of a sequence $\{\psi_j\}$ of automorphisms belonging to A_k. If j is sufficiently large, we have $\psi_j = \exp(L_j)\,\psi$ with $L_j \in \mathfrak{L}$ by the preceding lemma, and since $\psi_j \in A_k$ it follows that $\exp(L)\,\psi$ has at most k eigenvalues for all L near L_j. Since $n(\psi) > k$, it follows that there exists an $L' \in \mathfrak{L}$ such that $\exp(L')\psi$ has more than k eigenvalues. But then $\exp((1-z)L_j + zL')\psi$ has no more than k eigenvalues for any small z, but has more than k eigenvalues for $z = 1$, which is impossible by Lemma 0.29. This contradiction shows that A_k is closed.

It follows that each of the sets $m(\psi) = k$ is both open and closed. Thus, the whole of A' belongs to a single set $m(\psi) = k_0$. It is plain from Definition 25 that ψ is regular if and only if ψ has exactly $R_0 = m(\psi)$ eigenvalues. Thus, no $\psi \in A'$ has more eigenvalues than ψ, and, if ψ has as many eigenvalues as ϕ, ψ is regular.

Next, let ϕ and ψ be two elements of $A'' \subseteq A''$. Let U be a neighborhood of the identity in A, to be specified below. Since A' is connected, we may find a sequence $\phi = \phi_1, \phi_2, \ldots, \phi_p = \psi$ of elements of A' such that $\phi_{i+1}\phi_i^{-1} \in U$, $i = 1, \ldots, p-1$. It is plain from Definition 25 that A'' is dense in A'; thus we may even suppose that $\phi_i \in A''$, $i = 1, \ldots, p$. If U is taken to be sufficiently small, then we may write $\phi_{i+1} = \exp(L_i)\phi_i$, with $L_i \in \mathfrak{L}$. Since ϕ_{i+1} and ϕ_i are regular, it follows by Lemma 0.27 and by fact that each of the ϕ_i has the same (maximum) number k_0 of eigenvalues that $\exp(zL_i)\phi_i$ has as many eigenvalues as ϕ_i both for all z near $z = 0$ and for all z near $z = 1$. But, by Lemma 0.29, the set Z_0 of complex numbers z for which $\exp(zL_i)\phi_i$ has fewer eigenvalues than ϕ_i consists of isolated points exclusively. Therefore, the complement Z_1 of Z_0 is connected; sinc by thee above $\exp(zL_i)\phi_i \in A''$ for all $z \in Z_1$, the present lemma follows immediately. Q.E.D.

30. Lemma. ("Diagonalization") Let ϕ and ψ be two regular automorphisms of the complex semi-simple Lie algebra $\hat{\Lambda}$. Suppose that ϕ and ψ belong to the same connected component A' of the group A of automorphisms of $\hat{\Lambda}$. Then there exists an inner automorphism τ of $\hat{\Lambda}$ such that ϕ and $\tau\psi\tau^{-1}$ commute totally.

Proof. Let A'' be the set of regular elements of A'; by the preceding lemma, A'' is connected. The relationship R between two automorphisms η_1 and η_2 defined by the statement "there exists an inner automorphism τ of $\hat{\Lambda}$ such that η_1 and $\tau\eta_2\tau^{-1}$ commute totally" is easily verified to be symmetric and transitive, and hence to be an equivalence relation. Thus, to prove the present lemma, it is sufficient to prove that ϕ and any $\psi \in A''$ sufficiently close to ϕ stand in the relation R.

Now, by Lemma 13, it follows that

$$\hat{\Lambda} = \{\lambda \in \Lambda \mid \phi(\lambda) = \lambda\} + \{\nu - \phi(\nu) \mid \nu \in \hat{\Lambda}\}.$$

By Lemma 0.36, the derivative of the expression

$$\phi(tv, s\lambda) = \exp(\mathrm{Ad}(tv)) \exp(\mathrm{Ad}(s\lambda)) \phi \exp(\mathrm{Ad}(-tv))$$

with respect to t at $t = s = 0$ is $\mathrm{Ad}(v)\phi - \phi\,\mathrm{Ad}(v) = \mathrm{Ad}(v - \phi(v))\phi$, and its derivative with respect to s at $t = s = 0$ is $\mathrm{Ad}(\lambda)\phi$.

On the other hand, arguing as in the first paragraph of the proof of the preceding lemma, we see that every element of A sufficiently close to the identity has the form $\exp(L)$, where L is such that $\exp(tL)$ is an automorphism of $\hat{\Lambda}$ for all small t. But then, differentiating the formula $\exp(tL)[\mu, \mu'] = [\exp(tL)\mu, \exp(tL)\mu']$ with respect to t and setting $t = 0$, we see that L satisfies the hypotheses of Lemma 10, so that $L = \mathrm{Ad}(\lambda_0)$ for some $\lambda_0 \in \hat{\Lambda}$. Thus, for all sufficiently small λ and μ and all t and s such that $|t| \leqslant 1$, $|s| \leqslant 1$, we may write

$$\phi(tv, s\lambda) = \exp(\mathrm{Ad}(\mu(tv, s\lambda)))\phi.$$

Our calculations above show that $\frac{\partial}{\partial t}\mu(tv, s\lambda)|_{t=s=0} = v$, and that $\frac{\partial}{\partial s}\mu(tv, s\lambda)|_{s=t=0} = \lambda$. It follows at once by the implicit function theorem, Lemma 0.37, that every automorphism sufficiently close to ϕ can be written as $\phi(v, \lambda)$ with $\lambda \in \hat{\Lambda}$, $\phi(\lambda) = \lambda$, and λ small. But $\phi(v, \lambda)$ is of the form $\tau \exp(\mathrm{Ad}(\lambda))\phi\tau^{-1}$, where τ is inner; and since $\phi(\lambda) = \lambda$, $\exp(\mathrm{Ad}(\lambda))\phi$ commutes with ϕ. Since λ is small, $\exp(\mathrm{Ad}(\lambda))\phi$ is close to ϕ. Thus, by Corollary 28, $\exp(\mathrm{Ad}(\lambda))\phi$ commutes totally with ϕ, and our proof is complete. Q.E.D.

31. Corollary. Every automorphism ϕ of a complex semi-simple Lie algebra $\hat{\Lambda}$ which is sufficiently close to the identity has the form $\phi = \exp(\mathrm{Ad}(\lambda))$, where $\lambda \in \hat{\Lambda}$.

The following simple lemmas and corollaries will facilitate the application of Lemma 30 in a variety of cases.

32. Lemma. Let $\hat{\Lambda}$ be a complex semi-simple Lie algebra, and let A be its group of automorphisms.

(a) If $\phi, \psi \in A$ then ϕ and ψ belong to the same connected component of A if and only if $\phi\psi^{-1}$ is inner.

(b) Let $\hat{M}$ be a Cartan subalgebra of $\hat{\Lambda}$, and $\{\alpha\}$ the collection of roots of $\hat{M}$ on $\hat{\Lambda}$. Then if $\mu \in \hat{M}$, the automorphism $\exp(\mathrm{Ad}(\mu))$ is regular if and only if none of the differences $\alpha(\mu) - \beta(\mu)$, where $\alpha, \beta \in \{\alpha\}$ and

$\alpha \neq \beta$, is of the form $2\pi iq$, with q integral. Such elements $\mu \in \hat{M}$, always exist and if an automorphism ϕ of $\hat{\Lambda}$ commutes with $\exp(\mathrm{Ad}(\mu))$, then ϕ is of the form $\exp(\mathrm{Ad}(\nu))$, with $\mu \in \hat{M}$.

Proof. (a) That $\phi\psi^{-1}$ must be inner if ϕ and ψ belong to the same component of A follows readily from Corollary 31. Conversely, if $\phi = \exp(\mathrm{Ad}(\lambda))\,\psi$ we may connect ϕ to ψ through the curve $\exp(t\,\mathrm{Ad}(\lambda))\psi$ in A, which makes it plain that ϕ and ψ belong to the same component of A if $\phi\psi^{-1}$ is inner.

(b) If $\mathrm{Ad}(\lambda_0)$ has as many eigenvalues as any transformation of the form $\mathrm{Ad}(\lambda)$, then so does each transformation $\mathrm{Ad}(\lambda_1)$ with λ_1 sufficiently cleso to λ_0. But then, by Lemma II.2.8, each λ_1 sufficiently close to λ_0 belongs to a Cartan subalgebra, and therefore, by Theorem II.5.1 and the remark which follows it, the transformation $\mathrm{Ad}(\lambda_1)$ is similar to a transformation $\mathrm{Ad}(\mu)$ with $\mu \in \hat{M}$: $\mathrm{Ad}(\mu) = \tau\,\mathrm{Ad}(\lambda_1)\,\tau^{-1}$, where τ is an inner automorphism. If we let λ_0 belong to a neighborhood U of $\lambda = 0$ sufficiently small so that $\lambda \to \exp(\mathrm{Ad}(\lambda))$ is a homeomorphism of U onto a neighborhood of the identity of A, and use the evident fact that the regular automorphisms are dense in A, we see that one of the transformations $\exp(\mathrm{Ad}(\lambda_1))$ will be regular. But then the similar transformation $\exp(\mathrm{Ad}(\mu))$ must also be regular, as the reader will verify at once from Definition 25. Thus, by (a) and by the first sentence of the second paragraph of the proof of Lemma 16, $\exp(\mathrm{Ad}(\mu'))$ is regular if and only if it has as many distinct eigenvalues as any transformation $\exp(\mathrm{Ad}(\mu))$, $\mu \in \hat{M}$. If $\mu' \in \hat{M}$, it follows immediately from Lemma II.3.7 that $\mathrm{Ad}(\mu)\,\lambda_\alpha = \alpha(\mu)\,\lambda_\alpha$ if λ_α belongs to the root space Λ_α, so that the eigenspaces of $\exp(\mathrm{Ad}(\mu))$ are sums of the root spaces Λ_α of $\hat{M}$ on $\hat{\Lambda}$, and so that the number of eigenvalues of $\exp(\mathrm{Ad}(\mu'))$ is simply the number of distinct quantities $\exp(\alpha(\mu'))$, $\alpha \in \{\alpha\}$. Now, $\mu' \in \hat{M}$ may be chosen so that all these quantities are distinct. Indeed, if we let Π be a simple set of roots, and choose μ' so that the quantities $\alpha(\mu')$, $\alpha \in \Pi$, satisfy no relation of linear dependence with rational coefficients, it follows by Definition II.3.17 that all the quantities $\exp(\alpha(\mu'))$ are distinct.

This proves all but the final assertion of (b). To prove this last assertion, let μ and ϕ be as in (b). We have noted in the preceding paragraph that the eigenspaces of $\exp(\mathrm{Ad}(\mu))$ are simply the root spaces Λ_α; since all the quantities $\exp(\alpha(\mu))$ are distinct, all these eigenspaces belong to distinct eigenvalues of $\exp(\alpha(\mu))$. Thus, since ϕ commutes with $\exp(\alpha(\mu))$, ϕ maps each eigenspace Λ_α onto itself. Therefore ϕ maps each of the vectors μ_α of Corollary II.3.8 onto itself; since by Corollary II.3.10 these vectors span $\hat{M}$, $\phi|\hat{M} = I$. The final assertion of (b) now follows at once from Lemma 5.11. Q.E.D.

33. Lemma. Let $\hat{\Lambda}$ be a complex semi-simple Lie algebra, and let ϕ be an automorphism of $\hat{\Lambda}$. Suppose that there exists a complete basis for $\hat{\Lambda}$ consisting of eigenvectors of ϕ. Then there exists a regular automorphism ψ of $\hat{\Lambda}$, such that $\phi\psi^{-1}$ is inner, which commutes with ϕ.

Proof: Arguing precisely as in the last three paragraphs of the proof of Lemma 30, we find that every automorphism in the neighborhood of ϕ can be written in the form

$$\exp(\mathrm{Ad}(\nu))\exp(\mathrm{Ad}(\lambda))\,\phi\,\exp(-\mathrm{Ad}(\nu)), \qquad [*]$$

where $\lambda, \nu \in \hat{\Lambda}$, and $\phi(\lambda) = \lambda$. Since we have remarked that the regular automorphisms are dense in the set of all automorphisms, it follows that some one of the automorphisms [*] is regular. But then it follows from Definition 25 the equivalent automorphism $\exp(\mathrm{Ad}(\lambda))\,\phi$ is regular also. Using the fact that $\phi(\lambda) = \lambda$, the reader will verify easily that ϕ and $\exp(\mathrm{Ad}(\lambda))\,\phi$ commute, completing the proof of the present lemma. Q.E.D.

34. Corollary. The automorphism ψ of the preceding lemma may be taken to have the form $\exp(\mathrm{Ad}(\lambda))\,\phi$ where $\phi(\lambda) = \lambda$.

35. Corollary. Let $\hat{\Lambda}$ be a complex semi-simple Lie algebra, and let ϕ be an automorphism of $\hat{\Lambda}$. Suppose that there exists a complete basis for $\hat{\Lambda}$ consisting of eigenvectors of ϕ. Let ϕ_1 be a regular automorphism of $\hat{\Lambda}$, and suppose that $\phi\phi_1^{-1}$ is inner. Then there exists an inner automorphism τ of $\hat{\Lambda}$ such that ϕ and $\tau\phi\tau^{-1}$ commute.

Proof. Let ψ be as in Lemma 33, and, using Lemma 32(a), let τ be as in Lemma 30. Q.E.D.

We are now in a position to begin our study of the conjugations in the five exceptional simple complex Lie algebras. We begin with the simplest case, the Lie algebra of type G_2.

A′. Analysis of the conjugations in the Lie algebra of type G_2

As we saw in Section 4 (whose notations and results we take over for the next few paragraphs) a model for the Lie algebra $\hat{\Lambda}$ of type G_2 is given as a direct sum $\Lambda \oplus V \oplus V'$ of the Lie algebra Λ of type A_2 consisting of all 3×3 matrices λ of trace zero, a 3-dimensional vector space V, and the

3-dimensional "dual" vector space V' of all linear functionals on V. The set $\hat{M}$ of all diagonal matrices in Λ is a Cartan subalgebra. As we saw in the course of the proof of Lemma 21, the natural bilinear form satisfies $(\lambda + v + v', \lambda + v + v') = a \operatorname{tr}(\lambda^2) - bv'(v)$ with $a > 0$ and $b < 0$; the distinguished conjugation C of Lemma 9 is $\lambda \to -\bar{\lambda}'$, and $v \to Tv = u'$, $v' \to Tv' = u$, where $[u_1', u_2', u_3'] = (2/27)^{-1/3}[\bar{v}_1, \bar{v}_2, \bar{v}_3]$, $[u_1, u_2, u_3] = (2/27)^{1/3}[\bar{v}_1', \bar{v}_2', \bar{v}_3']$. By Lemma 12, and since the Dynkin diagram has no symmetries, every automorphism is inner. By Corollary 16, every conjugation C_1 in $\hat{\Lambda}$ is equivalent to a conjugation of the form ϕC, where ϕ is an automorphism, $\phi^2 = \phi$, and $\phi_c = \phi$. By Lemma 32 and Corollary 35, there exists an automorphism of the form $\psi = \exp(\operatorname{Ad}(\nu))$, where $\nu \in \hat{M}$, and an automorphism τ such that $\tau\phi\tau^{-1} = \exp(\operatorname{Ad}(\nu))$. Thus, $\exp(\operatorname{Ad}(\nu))^2 = 1$. If we put $d = \exp(\nu)$, so that d is a diagonal matrix, we find that $d^2 = I$, so that the entries of d are ± 1. Therefore $\psi_c = \psi$, and, by Corollary 18, the conjugations ϕC and ψC are equivalent. Summing up, we see that every conjugation is equivalent to a conjugation of the form $\exp(\operatorname{Ad}(\mu))C$, where $\mu \in \hat{M}$ and where μ is such that $\exp(\operatorname{Ad}(\mu))^2 = I$. If we write $\exp(\mu) = d$ as above, we have $\exp(\operatorname{Ad}(\mu))\, v = dv$ for each $v \in V$; thus we have $d^2 = \pm 1$. Since $\det(d) = 1$, we may suppose on renumbering coordinates that the diagonal entries of d are either $(1, 1, 1)$ or $(1, -1, -1)$. Calling the second diagonal matrix I^-, we find two real Lie algebras: the set $HG_2^{(3)}$ of all $\hat{\lambda} \in \hat{\Lambda}$ satisfying $C(\hat{\lambda}) = \hat{\lambda}$, and the set $HG_3^{(1)}$ of all $\hat{\lambda} \in \hat{\Lambda}$ satisfying $I^- C(\hat{\lambda})\, I^- = \hat{\lambda}$. Using the expression for the natural bilinear form of $\hat{\Lambda}$ given above, the reader will easily verify that the signature of $HG_2^{(3)}$ is zero, and that the signature of $HG_3^{(1)}$ is 8.

B'. Analysis of the conjugations in the Lie algebra of type E_8

A model for the Lie algebra $\hat{\Lambda}$ of type E_8 is given as a direct sum $\Lambda \oplus V \oplus V'$ of the Lie algebra of type A_8, consisting of all 9×9 matrices of trace zero, the vector space V of all tensors v_{ijk} of three indices, skew symmetric in the three indices, and the "dual" vector space V' of all tensors $(v')^{ijk}$ of three indices, skew-symmetric in the three indices. The set $\hat{M}$ of all diagonal matrices in Λ is a Cartan subalgebra. As we saw in the course of the proof of Lemma 9, the natural bilinear form satisfies $(\lambda + v + v', \lambda + v + v') = a \operatorname{tr}(\lambda^2) - bv'(v)$ with $a > 0$ and $b < 0$; the distinguished conjugation C of Lemma 9 is $\lambda \to -\bar{\lambda}'$, and $v \to Tv = u'$, $v' \to Tv' = u$, where $(u')^{ijk} = 2^{-5/3}3^{-5/3}7^{-1/3}\bar{v}_{ijk}$, $u_{ijk} = 2^{7/3}3^{7/3}7^{1/3}(\bar{v}')^{ijk}$. By Lemma 12, and since the Dynkin diagram has no symmetries, every automorphism is inner.

The set Π consisting of the roots $\omega_1 - \omega_2, \omega_2 - \omega_3, \ldots, \omega_7 - \omega_8$, $\omega_6 + \omega_7 + \omega_8$ is a simple set of roots. By Corollary 16, every conjugation C_1 in $\hat{\Lambda}$ is equivalent to a conjugation of the form ϕC, where ϕ is an automorphism, $\phi^2 = \phi$, and $\phi_c = \phi$. By Lemma 32 and Corollary 35, there exists an automorphism of the form $\psi = \exp(\mathrm{Ad}(\nu))$, where $\nu \in \hat{M}$, and an automorphism τ, such that $\tau\phi\tau^{-1} = \exp(\mathrm{Ad}(\nu))$. Thus $\psi^2 = I$. It follows that $\exp(\alpha(\nu))\,\lambda_\alpha = \psi\lambda_\alpha = \pm\lambda_\alpha$ for each root α of $\hat{M}$ in $\hat{\Lambda}$ and each λ_α in the corresponding root space Λ_α; thus $\exp(\alpha(\nu)) = \pm 1$. From the above description of C, and from the description of the root spaces Λ_α given in Section II.4, we see that $C\lambda_\alpha \in \Lambda_{-\alpha}$ for each root α. Thus $C\psi C = \psi$, and it follows by Corollary 18 that the conjugations ϕC and ψC are equivalent.

By Corollary II.3.19(iv), the vectors $\lambda_{\pm\alpha}, \alpha \in \Pi$, generate $\hat{\Lambda}$. Examining the action of the automorphism ψ on the vectors, and using the description of these vectors given in Section II.4, we see that there exists a diagonal matrix d with entries ± 1, such that ψ has the form $\lambda \to d\lambda d^{-1}$, $v_{ijk} \to d_i d_j d_k v_{ijk}$, and $(v')^{ijk} \to d_i^{-1} d_j^{-1} d_k^{-1} (v')^{ijk}$; the entry $d_9 = \pm 1$ is determined by the condition $\det(d) = 1$, which makes the preceding transformation an (inner) automorphism. Thus we see that every conjugation in $\hat{\Lambda}$ is equivalent to a conjugation ψC, where ψ is as above. We may suppose on renumbering coordinates that d is the diagonal matrix I_j^- of $9 - j$ 1's followed by j -1's, where $j = 0, 2, 4, 6$, or 8. Using the expression for the natural bilinear form of $\hat{\Lambda}$ given above, the reader will verify easily that the signature of the real Lie algebra determined by the condition $\hat{\lambda} = I_j^- C(\lambda)\, I_j^-$ is $2j((9-j) + \dfrac{(j-1)(j-2)}{6} + \dfrac{(9-j)(8-j)}{2}$, so that for $j = 0, 2, 4, 6, 8$ we find the signatures 0, 112, 128, 112, 128. This suggests that two pairs of our five Lie algebras are isomorphic, a conjecture which is in fact true, and which we may verify by the following calculations.

According to the relevant passage in Section 4, the roots of $\hat{M}$ in $\hat{\Lambda}$ are $\pm(\omega_i - \omega_j)$ where $j < j$, and $\pm(\omega_i + \omega_j + \omega_k)$ where $i < j < k$; here $1 \leqslant i, j, k \leqslant 9$, and $\omega_1 + \cdots + \omega_9 = 0$. Let α be one of the roots $\omega_i + \omega_j + \omega_k$, and $v = v^{(i,j,k)}$ (resp. $v' = v'_{(i,j,k)}$) denote the skew-symmetric tensor whose element $v^{(i,j,k)}_{l,m,n}$ (resp. $(v'_{(i,j,k)})^{l,m,n}$) is zero unless l, m, n and i, j, k constitute the same set of 3 elements, and which is equal to the sign of the permutation $\binom{ijk}{lmn}$ if l, m, n and i, j, k constitute the same set of 3 elements. Then $v^{(i,j,k)}$ belongs to the root space Λ_α and $v'_{(i,j,k)}$ to the root space $\Lambda_{-\alpha}$. We have $(v', v) = \frac{1}{2} b v'(v) < 0$; thus we can choose real numbers $r_1 > 0$ and $r_2 < 0$ such that $\lambda_\alpha = r_1 v$ and $\lambda_{-\alpha} = r_2 v$ satisfy $(\lambda_\alpha, \lambda_{-\alpha}) = 1$ and $C\lambda_\alpha = -\lambda_{-\alpha}$, $C\lambda_{-\alpha} = -\lambda_\alpha$. Now, we saw in the course of the proof of Lemma II.5.9 that the automorphism ϕ_α defined by

$$\phi_\alpha = \exp\left(i\pi(2\alpha(\mu_\alpha))^{-1/2}\operatorname{Ad}(\lambda_\alpha + \lambda_{-\alpha})\right)$$

maps $\hat{M}$ into $\hat{M}$, and that $\phi_\alpha(\mu_\beta) = \mu_{W_\alpha\beta}$, where W_α is the Weyl reflection of Definition II.5.8. Moreover, we may deduce from the formulae for

$$\exp\left(it(\lambda_\alpha + \lambda_{-\alpha})\right)\mu_\alpha \quad \text{and} \quad \exp\left(it(\lambda_\alpha + \lambda_{-\alpha})\right)(\lambda_\alpha - \lambda_{-\alpha})$$

given in that proof that $\phi_\alpha^2\lambda_{\pm\alpha} = \lambda_{\pm\alpha}$. It follows from the above that $C\phi_\alpha = \phi_\alpha C$. Thus, if $\nu \in \hat{M}$ and $\exp(\operatorname{Ad}(\nu))\,\lambda_{\pm\alpha} = -\lambda_{\pm\alpha}$, we have

$$\phi_\alpha \exp(\operatorname{Ad}(\nu))\,C\phi_\alpha^{-1} = \phi_\alpha^2 \exp(\operatorname{Ad}(\nu))\,C.$$

Suppose that β is a root of the form $\omega_k - \omega_l$ such that $\beta + \alpha$ is a root, so that, examining the full set of roots, we see that neither $\beta - \alpha$ nor $\beta + 2\alpha$ is a root. Choose λ_β in the root space Λ_β, and put $\lambda_{\alpha+\beta} = [\lambda_\alpha, \lambda_\beta]$ so that $\lambda_{\alpha+\beta} \neq 0$ by Lemma 3.11, and so that $\operatorname{Ad}(\lambda_\alpha + \lambda_{-\alpha})\,\lambda_{\alpha+\beta} = -[\lambda_\beta, [\lambda_{-\alpha}, \lambda_\alpha]] = -\beta(\mu_\alpha)\,\lambda_\beta = \frac{1}{2}\alpha(\mu_\alpha)\,\lambda_\beta$ by Lemmas II.2.5 and II.3.7. Introducing $\lambda'_\beta = (\frac{1}{2}\alpha(\mu_\alpha))^{1/2}\,\lambda_\beta$ and $\lambda'_{\alpha+\beta} = \lambda_{\alpha+\beta}$, we have

$$\operatorname{Ad}(\lambda_\alpha + \lambda_{-\alpha})\,\lambda'_\beta = (2\alpha(\mu_\alpha))^{1/2}\cdot\tfrac{1}{2}\lambda'_{\alpha+\beta}$$

$$\operatorname{Ad}(\lambda_\alpha + \lambda_{-\alpha})\,\lambda'_{\beta+\alpha} = (2\alpha(\mu_\alpha))^{1/2}\cdot\tfrac{1}{2}\lambda'_\beta.$$

Thus, by the exponential series, $\phi_\alpha^2\lambda'_\beta = -\lambda'_\beta$. On the other hand, if β is a root which is not a neighbor of α, then plainly $\phi_\alpha^2\lambda_\beta = \lambda_\beta$. Since $\phi_\alpha\mu_\gamma = \mu_{W_\alpha\gamma}$ for each root γ, and since $W_\alpha^2 = I$ by Definition II.5.8, we have $\phi_\alpha^2 = \exp(\operatorname{Ad}(\eta))$ by Lemma II.5.1, where $\eta \in \hat{M}$. Putting $\delta = \exp(\eta)$, δ is a diagonal matrix such that $\delta^2\lambda_\alpha\delta^{-2} = \lambda_\alpha$ for each α of the form $\omega_i - \omega_j$; thus $\delta^2\lambda\delta^{-2} = \lambda$ for all $\lambda \in \hat{\Lambda}$, so that $\delta^2 = a_1 I$. On the other hand, $\delta v = v$ if v is a vector in a root space Λ_β, where $\beta = \omega_k + \omega_l + \omega_m$ is not a neighbor of α. Therefore $a_1^3 = 1$, and δ is a diagonal matrix δ' of zeroes and ones, multiplied by a cube root of unity. But then we verify easily that the automorphism ϕ_α^2 has the form $\lambda \to \delta'\lambda\delta'$, $v_{pqr} \to \delta'_p\delta'_q\delta'_r v_{pqr}$, etc. Since $\det(\delta) = 1$, while by the above $\delta e_{pq}\delta^{-1} = \exp(\operatorname{Ad}(\eta))\,e_{pq} = \pm e_{pq}$, the plus sign to be applied if p and q either both do or both do not belong to the set i, j, k, and the minus sign to be applied otherwise, we see that δ must be the diagonal matrix of entries δ_p, where $\delta_p = -1$ if $p \neq i, j, k$, while $\delta_p = 1$ if $p = i, j$, or k. Thus, taking all together, we see that the conjugations $\lambda \to dC(\lambda)\,d^{-1}$, etc., and $\lambda \to \delta dC(\lambda)\,d^{-1}\delta^{-1}$, etc., are equivalent if $d\lambda_{\pm\alpha} = -\lambda_{\pm\alpha}$.

Remark: We will need the result of calculations exactly like the preceding calculation repeatedly in what is to follow, and, in order to avoid unnecessary repetition, we now summarize its essential features here for future reference. Suppose that $\tilde{\Lambda}$ is a semi-simple complex Lie algebra, and that $\tilde{M}$ is a Cartan subalgebra of $\tilde{\Lambda}$. Let $\{\alpha\}$ be the set of roots of $\tilde{M}$ in $\tilde{\Lambda}$, and Λ_α the corresponding root spaces. Let $\alpha \in \{\alpha\}$ be non-zero, let $\lambda_\alpha \in \Lambda_{\pm\alpha}$, let $(\lambda_\alpha, \lambda_{-\alpha}) = 1$, and let ϕ_α be the automorphism of $\tilde{\Lambda}$ defined by

$$\phi_\alpha = \exp\,(i\pi(2\alpha(\mu_\alpha))^{-1/2}\,\mathrm{Ad}\,(\lambda_\alpha + \lambda_{-\alpha})).$$

Then the above calculations show that $\phi_\alpha^2 = \exp\,(\mathrm{Ad}\,(\eta))$, where $\eta \in \tilde{M}$. Moreover, $\exp\,(\alpha(\eta)) = 1$, $\exp\,(\beta(\eta)) = 1$ if $0 \neq \beta \in \{\alpha\}$ is not a neighbor of α, and $\exp\,(\beta(\eta)) = -1$ if $0 \neq \beta \in \{\alpha\}$ is a simple neighbor of α. If C is a positive conjugation in $\tilde{\Lambda}$ such that $C\tilde{M} = \tilde{M}$ and such that $C\Lambda_\alpha = \Lambda_{-\alpha}$ for each $\alpha \in \{\alpha\}$, then it follows as above that we may choose λ_α and $\lambda_{-\alpha}$ so that $C\lambda_\alpha = -\lambda_{-\alpha}$; in this case we have $C\phi_\alpha = \phi_\alpha C$. Thus, if we have a conjugation of the form τC, where the automorphism τ satisfies $\tau(\lambda_{\pm\alpha}) = -\lambda_{\pm\alpha}$, we find, calculating as above, that τC and $\phi_\alpha^2 \tau C$ are equivalent conjugations.

If $d = I_4^-$, and $i, j, k = 7, 8, 9$, δd is a diagonal matrix of eight -1's and one $+1$; if $d = I_6^-$ and $i, j, k = 1, 2, 9$, δd is a diagonal matrix of two -1's and seven $+1$'s. Then, renumbering coordinates, we see that the conjugations $\lambda \to I_6^- C(\lambda)\, J_6^-$, etc., $\lambda \to I_2^-\ C(\lambda)\, I_2^-$, etc. are equivalent, as are the conjugations $\lambda \to I_4 C(\lambda)\, I_4^-$, etc., and $\lambda \to I_8^- C(\lambda)\, I_8^-$, etc. We therefore find three real Lie algebras: the algebra $HE_8^{(9)}$ of all $\hat{\lambda} \in \hat{\Lambda}$ satisfying $\hat{\lambda} = C(\hat{\lambda})$, which has signature 0; the algebra $HE_8^{(7)}$ of all $\hat{\lambda} \in \hat{\Lambda}$ satisfying $\hat{\lambda} = I_2^- C\,(\tilde{\lambda})\, I_2^-$, etc., which has signature 112, and the algebra $HE_8^{(5)}$ of all $\hat{\lambda} \in \hat{\Lambda}$ satisfying $\hat{\lambda} = I_4^- C\,(\hat{\lambda})\, I_4^-$, etc. which has signature 128.

C′. Analysis of the conjugations in the Lie algebra of type E_7

Let $\hat{\Lambda}$ be the Lie algebra of type E_8 discussed just above, and let C be the positive conjugation in $\hat{\Lambda}$ to which reference is made above. As we saw in Section II.4 (whose notations and results we use freely in what follows) a model for the Lie algebra of type E_7 is furnished by the subalgebra $\hat{\Lambda}^{(0)}$ of $\hat{\Lambda}$ spanned by all the elements μ_α and $\lambda_{\pm\alpha}$ for which the root α is linearly dependent on the roots belonging to the simple set $\Pi^{(0)} = \{\omega_2 - \omega_3, \ldots, \omega_7 - \omega_8, \omega_6 + \omega_7 + \omega_8\}$ (cf. the relevant passages in Section II.4). As we saw in Section II.4, $\hat{\Lambda}^{(0)}$ is the direct sum $\Lambda^{(0)} = \Lambda^{(0)} \oplus V^{(0)} \oplus V'^{(0)}$, where $\Lambda^{(0)}$ is the set of all 9×9 matrices $\lambda = (\lambda_{ij})$ of trace zero such that $\lambda_{j9} = \lambda_{9j} = 0$ if $j \neq 9$, $\lambda_{j1} = \lambda_{1j} = 0$ if $j \neq 1$, and $\lambda_{11} = \lambda_{99}$; $V^{(0)}$ is the set of skew-symmetric tensors with components v_{ijk} which vanish whenever just one of i, j, k is 1 or 9; and $V'^{(0)}$ is the dual vector space of all skew-symmetric tensors with components $(v')^{ijk}$ which vanish whenever just one of i, j, k is 1 or 9. The distinguished conjugation of Lemma 9 is the restriction to $\hat{\Lambda}^{(0)}$ of C; as we saw in proving Lemma 9, the natural bilinear form in $\hat{\Lambda}^{(0)}$ is a positive multiple of the restriction to $\hat{\Lambda}^{(0)}$ of the natural bilinear form in $\hat{\Lambda}$, and thus satisfies $(\lambda + v + v', \lambda + v + v') = a \operatorname{tr}(\lambda^2) - bv'(v)$, where $a > 0$, $b < 0$. The set $\hat{M}^{(0)}$ of all diagonal matrices in $\Lambda^{(0)}$ is a Cartan subalgebras of $\hat{\Lambda}^{(0)}$. By Lemma 12, and since the Dynkin diagram has no symmetries, every automorphism is inner.

Arguing precisely as in the second paragraph under heading B', we verify the following statement: Let C_1 be a conjugation in $\hat{\Lambda}^{(0)}$. There exists a diagonal matrix d with entries d_j, such that $d_j = \pm 1$ for $2 \leqslant j \leqslant 8$, such that $d_1 = d_9 = \pm 1$ or $d_1 = d_8 = \pm i$, and such that $\det(d) = 1$, so that the transformation ψ: $\lambda \to d\lambda d^{-1}$, $v_{klm} \to d_k d_l d_m v_{klm}$, $(v')^{klm} \to d_k^{-1} d_l^{-1} d_m^{-1} (v')^{klm}$ is an inner automorphism of $\hat{\Lambda}$ mapping $\hat{\Lambda}^{(0)}$ into itself. This diagonal matrix d may be chosen so that the conjugation ψC is equivalent to the conjugation C_1.

Note that d_1 and d_9 are determined up to sign by $d_2, \ldots, d_8$ and by the conditions $d_1 = d_8$ and $\det(d) = 1$; however, we may change the sign of d_1 and d_9 without affecting the restriction of the above-described automorphism ψ to the subalgebra $\hat{\Lambda}^{(0)}$ of $\hat{\Lambda}$. Since it is always possible to renumber the entries $d_2, \ldots, d_8$ of d, the equivalence class of the conjugation ψC is determined simply by the number n_+ of $+1$'s among $d_2, \ldots, d_8$. Thus we must deal with 8 separate conjugations. However, by the remark made in the semi-final paragraph under the preceding heading, we may, on passing to an equivalent conjugation, replace d by δd: here, the diagonal matrix δ has diagonal entries δ_j which are equal to 1 if j is one of three integers r, s, t, and equal to -1 otherwise, and we require that $d_r d_s d_t = -1$. Note also that in order to insure that the automorphism $\exp(\operatorname{Ad}(\lambda + \lambda'))$ of the cited remark maps $\hat{\Lambda}^0$ into itself, we must require that either none or two of r, s, t shall belong to the subset $\{1, 9\}$ of $\{1, 2, \ldots, 9\}$. Using this remark, the reader will easily verify that of our 8 cases $n_+ = 0, \ldots, 7$, the cases $n_+ = 1$, 3, and 5 correspond to equivalent conjugations, the cases $n_+ = 0$ and $n_+ = 4$ correspond to equivalent conjugations, and the cases $n_+ = 1$ and $n_+ = 6$ correspond to equivalent conjugations. Thus we have only to consider the four cases $n_+ = 4, 5, 6, 7$, and therefore we may take the matrix d to have one of the four following sets of diagonal entries:

$$d^{(7)}\colon (1, 1, 1, 1, 1, 1, 1, 1, 1) \text{ (so that } d^{(7)} = I);$$

$$d^{(6)}\colon (i, 1, 1, 1, 1, 1, 1, -1, i)$$

$$d^{(5)}\colon (1, 1, 1, 1, 1, 1, -1, -1, 1)$$

$$d^{(4)}\colon (i, 1, 1, 1, 1, -1, -1, -1, i).$$

Thus we find the four real Lie algebras $HE_7^{(7)}$, $HE_7^{(6)}$, $HE_7^{(5)}$, and $HE_7^{(4)}$ determined by the conditions $\lambda = d^{(j)}C(\lambda)(d^{(j)})^{-1}$ etc., $j = 4, 5, 6, 7$. Using the description of $\hat{\Lambda}^{(0)}$ and its natural bilinear form given in the first paragraph under the present heading C', we readily calculate the signatures of these four real simple Lie algebras to be 0, 54, 64, and 70, so that no two of them are isomorphic.

D'. Analysis of the conjugations in the Lie algebra of type E_6

Let $\hat{\Lambda}$ be the Lie algebra of type E_8 discussed above under heading B', and let C be a positive conjugation in $\hat{\Lambda}$ as under heading B'. As we saw in Section II.4 (whose notations and results we use freely in what follows) a model for the Lie algebra of type E_6 is furnished by the subalgebra $\hat{\Lambda}^{(1)}$ of $\hat{\Lambda}$ spanned by the elements μ_α and $\lambda_{\pm\alpha}$ for which the root α is linearly dependent on the roots belonging to the simple set $\Pi^{(1)} = \{\omega_3 - \omega_4, \ldots, \omega_7 - \omega_8, \omega_6 + \omega_7 + \omega_8\}$ (cf. the relevant passages in Section II.4). As we saw in Section II.4, $\hat{\Lambda}^{(1)}$ is the direct sum $\hat{\Lambda}^{(1)} = \Lambda^{(1)} \oplus V^{(1)} \oplus V'^{(1)}$, where $\Lambda^{(1)}$ is the set of all 9×9 matrices $\lambda = (\lambda_{ij})$ of trace zero such that $\lambda_{j9} = \lambda_{9j} = 0$ if $j \neq 9$, $\lambda_{j1} = \lambda_{1j} = 0$ if $j \neq 1$, and $\lambda_{2j} = \lambda_{j2} = 0$ if $j \neq 2$, while $\lambda_{11} = \lambda_{22} = \lambda_{99}$; $V^{(1)}$ is the set of skew-symmetric tensors v with components v_{ijk} which vanish whenever one or two but not all three of i, j, k are 1, 2, or 9; and $V'^{(1)}$ is the dual vector space of all skew-symmetric tensors with components $(v')^{ijk}$ which vanish whenever one or two but not all three of i, j, k is 1, 2, or 9. The distinguished conjugation of Lemma 9 is the restriction to $\hat{\Lambda}^{(1)}$ of the conjugation C of $\hat{\Lambda}$; as we saw in proving Lemma 9, the natural bilinear form in $\hat{\Lambda}^{(1)}$ is a positive multiple of the restriction to $\Lambda^{(1)}$ of the natural bilinear form in $\hat{\Lambda}$, and thus satisfies $(\lambda + v + v', \lambda + v + v') = a \operatorname{tr}(\lambda^2) - bv'(v)$,where $a > 0$, $b < 0$. The set $\hat{M}^{(1)}$ of all diagonal matrices in $\Lambda^{(1)}$ is a Cartan subalgebra of $\hat{\Lambda}^{(1)}$. The roots of $\hat{M}^{(1)}$ in $\hat{\Lambda}^{(1)}$ are $\omega_i - \omega_j$, where $i \neq j$ and $3 \leqslant i, j \leqslant 8$; $\pm(\omega_i + \omega_j + \omega_k)$, where i, j, k are all distinct and $3 \leqslant i, j, k \leqslant 8$; and $\pm(\omega_3 + \cdots + \omega_8)$. Noting the symmetries of the Dynkin diagram, and using Lemma 12, we see that if A is the group of all automorphisms of $\hat{\Lambda}^{(1)}$, and A_0 the normal subgroup consisting of all inner automorphisms, then A/A_0 is of order at most 2. We see in fact, examining the proof of Lemma 12, that if we number the roots $\omega_3 - \omega_4, \omega_4 - \omega_5, \omega_5 - \omega_6, \omega_6 - \omega_7, \omega_7 - \omega_8, \omega_6 + \omega_7 + \omega_8$ as $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha'$, choose $\lambda_{\pm\alpha_i}$ and $\lambda_{\pm\alpha'}$ in the root spaces $\Lambda_{\pm\alpha_i}$ and $\Lambda_{\pm\alpha'}$ respectively in such a way that $(\lambda_{\alpha_i}, \lambda_{-\alpha_i}) = (\lambda_{\alpha'}, \lambda_{-\alpha'}) = 1$, and let Φ be the automorphism of $\hat{\Lambda}^{(1)}$ defined by

$$\Phi\colon \lambda_{\pm\alpha_1}, \lambda_{\pm\alpha_2}, \lambda_{\pm\alpha_3}, \lambda_{\pm\alpha_4}, \lambda_{\pm\alpha_5}, \lambda_{\pm\alpha'}$$

$$\to \lambda_{\pm\alpha_5}, \lambda_{\pm\alpha_4}, \lambda_{\pm\alpha_3}, \lambda_{\pm\alpha_2}, \lambda_{\pm\alpha_1}, \lambda_{\pm\alpha'}$$

(cf. Lemma II.3.10 and its proof), then every automorphism of $\hat{\Lambda}^{(1(}$ is either inner, or is the product of Φ and an inner automorphism. Note that $\Phi^2 = I$. For later purposes we wish to be little more specific about the choice of the vectors $\lambda_{\pm\alpha}$ defining the automorphism Φ. We have noted above, in the course of the proof of Lemma 9, that $C\Lambda_\alpha = \Lambda_{-\alpha}$ for each root α. Thus, if $\lambda \in \Lambda_\alpha$, $C\lambda \in \Lambda_{-\alpha}$. We have $(\lambda, C\lambda) < 0$; thus, multiplying λ by a suitable constant, we may suppose that $(\lambda, C\lambda) = -1$. This shows that the vectors $\lambda_{\pm\alpha_i}$ and $\lambda_{\pm\alpha'}$ above may be chosen so that $\lambda_{-\alpha_i} = -C\lambda_{\alpha_i}$, etc. Then $C\Phi C\lambda_{\alpha_1} = -C\Phi\lambda_{-\alpha_1} = -C\lambda_{-\alpha_5} = \lambda_{\alpha_5} = \Phi\lambda_{\alpha_1}$, etc., so that $C\Phi C = \Phi$, i.e., $\Phi = \Phi_c$. Of course, we are still not sure that the automorphism Φ is not actually an inner automorphism.

Leaving this matter in abeyance for the moment however, we proceed to analyze the conjugations in $\hat{\Lambda}^{(1)}$ which are of the form τC, where the automorphism τ is inner. For this class of conjugations, we may argue precisely as in the second paragraph under heading B' above, to verify the following statement: Let C_1 be a conjugation in $\hat{\Lambda}^{(1)}$ of the form τC, where the automorphism is inner. There exists a diagonal matrix d with entries d_j all equal to ± 1, such that $d_1 = d_2 = d_9$ and $\det(d) = 1$, so that the transformation ψ: $\lambda \to d\lambda d^{-1}$, $v_{klm} \to d_k d_l d_m v_{klm}$, $(v')^{klm} \to d_k^{-1} d_l^{-1} d_m^{-1} (v')^{klm}$ is an inner automorphism of $\hat{\Lambda}$ mapping $\hat{\Lambda}^{(1)}$ into itself. This diagonal matrix d may be chosen so that the conjugation ψC is equivalent to the conjugation C_1. (Note that d_1, d_2, d_9 are determined by $d_3, \ldots, d_8$ and the conditions $d_1 = d_2 = d_9$ and $\det(d) = 1$.)

Since it is always possible to renumber the entries $d_3, \ldots, d_8$ of d, the equivalence class of the conjugation ψC is determined simply by the number n_+ of $+1$'s among $d_3, \ldots, d_8$. Thus we must deal with 7 separate conjugations. However, by the remark made in the semi-final paragraph under heading B' above, we may, on passing to an equivalent conjugation, replace d by δd: here, the diagonal matrix δ has diagonal entries δ_j which are equal to 1 if j is one of three integers r, s, t, and equal to -1 otherwise, and we require that $d_r d_s d_t = -1$. Note also that in order to insure that the automorphism $\exp(\mathrm{Ad}(\lambda + \lambda'))$ of the cited remark maps $\hat{\Lambda}^{(1)}$ into itself, we must require that $3 \leqslant r, s, t \leqslant 8$. Using all this, the reader will easily verify that of our 7 cases $n_+ = 0, \ldots, 6$, the cases $n_+ = 0, 3$, and 4 correspond to equivalent conjugations, and the cases $n_+ = 1, 2$, and 5 correspond to equivalent conjugations. Thus we have only to consider the four cases $n_+ = 4, 5, 6$, and therefore we may take the matrix d to have one of the three following sets of diagonal entries:

$$d^{(6)}: (1, 1, 1, 1, 1, 1, 1, 1, 1) \text{ (so that } d^{(6)} = I)$$

$$d^{(5)}: (-1, -1, 1, 1, 1, 1, 1, -1, -1)$$

$$d^{(4)}: (1, 1, 1, 1, 1, 1, -1, -1, 1).$$

We find the three real Lie algebras $HE_6^{(6)}$, $HE_6^{(5)}$, and $HE_6^{(4)}$ determined by the conditions $\lambda = d^{(j)} C(\lambda) (d^{(j)})^{-1}$, etc., $j = 4, 5, 6$. Using the description of $\hat{\Lambda}^{(1)}$ and its natural bilinear form given in the first paragraph under the present heading D', we readily calculate the signatures of these three real simple Lie algebras to be 0, 32, and 40.

It is now easy to show that the automorphism Φ is not inner. For if it were, all automorphisms of $\hat{\Lambda}^{(1)}$ would be inner, and every conjugation in $\hat{\Lambda}^{(1)}$ would be equivalent to one of the three conjugations described in the preceding paragraph. But $\hat{\Lambda}^{(1)}$ admits the conjugation $\lambda \to \bar{\lambda}$, $v \to \bar{v}$, $v' \to \bar{v}'$, and the signature of the real Lie algebra RE_6 defined by this conjugation is readily calculated to be 42. Thus A_0 has two cosets in A: A_0 itself, and the outer automorphisms $A_0 \Phi$, so that we must give a separate analysis of the conjugations of the form τC, where τ is outer. (Note that here we do not exploit the symmetry of the Dynkin diagram: rather, the symmetry of the Dynkin diagram exploits us.)

This we do as follows: choose $\mu \in \hat{M}^{(1)}$ such that the six numbers $\alpha_1(\mu), \alpha_2(\mu), \ldots, \alpha_5(\mu), \alpha'(\mu)$ satisfy no relation $n_1\alpha_1(\mu) + \cdots + n_6\alpha'(\mu) = 2\pi i n_7$ with integer coefficients $n_1, \ldots, n_7$, so that all the quantities $\exp(n_1\alpha_1(\mu) + \cdots + n_6\alpha'(\mu))$ are distinct. We have $(\exp(\mathrm{Ad}(\mu))\Phi)^2 = \exp(\mathrm{Ad}(\mu + \Phi\mu))$; since there is a complete basis for $\hat{\Lambda}^{(1)}$ consisting of eigenvectors of $\exp(\mathrm{Ad}(\mu + \Phi\mu))$, it follows that $(\exp(\mathrm{Ad}(\mu))\Phi)^2$ satisfies a polynomial equation without repeated roots, and therefore $\exp(\mathrm{Ad}(\mu))\Phi$ satisfies such a polynomial equation also, so that there is a complete basis for $\hat{\Lambda}^{(1)}$ consisting of eigenvalues of $\exp(\mathrm{Ad}(\mu))\Phi$. We now see by Corollary 34 that there exists a regular automorphism of the form $\exp(\mathrm{Ad}(\nu))\exp(\mathrm{Ad}(\mu))\Phi$, where $(\exp(\mathrm{Ad}(\mu))\Phi)(\nu) = \nu$. From the last equation we find $(\exp(\mathrm{Ad}(\mu))\Phi)^2(\nu) = \nu$; thus $\exp(\mathrm{Ad}(\mu + \Phi\mu))(\nu) = \nu$. Since $\Phi\mu_{\alpha_1} = \mu_{\alpha_5}$, etc., we have $\alpha_1(\Phi(\mu)) = \alpha_5(\mu)$, etc., by Corollary II.3.8; thus since $\mathrm{Ad}(\mu + \Phi\mu)\lambda_\alpha = \alpha(\mu + \Phi\mu)$ if λ_α belongs to the root space Λ_α, and by the way we have chosen μ, every root space Λ_α except $\hat{M}^{(1)} = \Lambda_0$ is contained in an eigenspace of $\exp(\mathrm{Ad}(\mu + \Phi\mu))$ belonging to an eigenvalue different from zero, and we can conclude that $\nu \in \hat{M}^{(1)}$. This shows that there exists a regular automorphism of $\hat{\Lambda}_1$ of the form $\exp(\mathrm{Ad}(\mu_1))\Phi$, where $\mu_1 \in \hat{M}^{(1)}$. The regular automorphisms form an open set by Lemma 0.27; thus if we vary μ_1 very slightly within $\hat{M}^{(1)}$, the automorphism $\exp(\mathrm{Ad}(\mu_1))\Phi$ remains regular. Thus we may assume without loss of generality that μ_1 has the same property as μ, that all the quantities $\exp(n_1\alpha_1(\mu_1) + \cdots + n_6\alpha'(\mu_1))$ are distinct; so that finally, replacing μ by μ_1, we see that, without loss of generality, we may suppose that $\mu_1 = \mu$, i.e., that $\exp(\mathrm{Ad}(\mu))\Phi$ is regular.

By Corollary 16, every conjugation of the form τC, where τ is an outer automorphism, is equivalent to a conjugation of the same form, where τ is an outer automorphism and satisfies $\tau^2 = I$ and $\tau_c = \tau$. By Lemma 32 and Corollary 35, there exists an inner automorphism σ of $\hat{\Lambda}^{(1)}$ such that $\tau_1 = \sigma\tau\sigma^{-1}$ and $\exp(\mathrm{Ad}(\mu))\Phi$ commute. Thus τ_1 maps each eigenspace of the automorphism $(\exp(\mathrm{Ad}(\mu))\Phi)^2 = \exp(\mathrm{Ad}(\mu + \Phi\mu))$ onto itself. We have seen above that, in virtue of the condition that all the quantities $\exp(n_1\alpha_1(\mu) + \cdots + n_6\alpha'(\mu))$ are distinct, it follows that the eigenspace of $\exp(\mathrm{Ad}(\mu + \Phi\mu))$ belonging to the eigenvalue 1 is simply the Cartan subalgebra $\hat{M}^{(1)}$. Thus τ_1 maps $\hat{M}^{(1)}$ onto itself, and consequently permutes the vectors μ_α of Corollary II.3.8 among themselves. Define $\Phi\alpha$ for each root α by the equation $\Phi\mu_\alpha = \mu_{\Phi\alpha}$. In virtue of the condition imposed on μ, it follows that if α and β are roots and $\exp(\alpha(\mu + \Phi\mu)) = \exp(\beta(\mu + \Phi\mu))$, then $\alpha + \Phi\alpha = \beta + \Phi\beta$. Thus the eigenspaces of $(\exp(\mathrm{Ad}(\mu))\Phi)^2$ are in the first place the space $\hat{M}^{(1)}$, and, in the second place, the various sums $\tilde{\Lambda}(\alpha_0) = \sum_{\alpha+\Phi\alpha=\alpha_0+\Phi\alpha_0} \Lambda_\alpha$ which can be formed from the root spaces Λ_α. On the other hand, as we noted in Section II.4, in the third paragraph following formula (12) of that section, $\alpha + \Phi\alpha = \beta + \Phi\beta$ implies that either $\alpha = \beta$ or $\alpha = \Phi\beta$. Thus, if $\alpha = \Phi\alpha$, Λ_α is invariant under τ_1; while if $\alpha \neq \Phi\alpha$, $\Lambda_\alpha + \Lambda_{\Phi\alpha}$ is invariant under τ_1. It follows that if $\alpha = \Phi\alpha$, then $\tau_1\mu_\alpha = \mu_\alpha$; if $\alpha \neq \Phi\alpha$, then by Lemma II.2.5 and Corollary II.3.8 and since $\hat{\Lambda}^{(1)}$ is the *direct* sum of the various root spaces, $[\Lambda_\alpha + \Lambda_{\Phi\alpha}, \Lambda_{-\alpha} + \Lambda_{-\Phi\alpha}] \cap \hat{M}^{(1)}$ is the space spanned by μ_α and $\mu_{\Phi\alpha}$, so that the space spanned by these two vectors is invariant under τ_1, and we have either $\tau_1\mu_\alpha = \mu_\alpha$ or $\tau_1\mu_\alpha = \mu_{\Phi\alpha}$. Therefore, in particular, τ_1 maps the set $\Pi^{(1)}$ of roots into itself. But since τ_1 is an automorphism, $\tau_1\alpha$ and $\tau_1\beta$ must be neighbors if α and β are neighbors (cf. Definition II.3.24). Thus the restriction of τ_1 to $\Pi^{(1)}$ must define a symmetry of the Dynkin diagram, so that either $\tau_1\mu_\alpha = \mu_\alpha, \alpha \in \Pi^{(1)}$, or $\tau_1\mu_\alpha = \Phi\mu_\alpha, \alpha \in \Pi^{(1)}$. But in the former case it would follow by Lemma II.5.11 that τ_1 were inner, which we know to be false. Hence $\tau_1\mu = \Phi\mu, \mu \in \hat{M}^{(1)}$, so that by Lemma II.5.11 there exists a $\hat{\mu} \in \hat{M}^{(1)}$ such that $\tau_1 = \exp(\mathrm{Ad}(\hat{\mu}))\,\Phi$.

We have

$$\exp(-\tfrac{1}{2}\mathrm{Ad}(\hat{\mu}))\exp(\mathrm{Ad}(\hat{\mu}))\,\Phi\exp(\mathrm{Ad}(\hat{\mu})) = \exp(\tfrac{1}{2}\mathrm{Ad}(\hat{\mu} + \Phi\hat{\mu}))\,\Phi.$$

Thus, putting $\hat{\nu} = \frac{1}{2}\hat{\mu} + \Phi\hat{\mu}$, we see that $\Phi\hat{\nu} = \hat{\nu}$, and that there exists an automorphism $\hat{\sigma}$ such that $\hat{\tau} = \hat{\sigma}\tau\hat{\sigma}^{-1} = \exp(\mathrm{Ad}(\hat{\nu}))\,\Phi$. Since $\tau^2 = I$, we have $\hat{\tau}^2 = \exp(\mathrm{Ad}(\hat{\nu}))^2 = I$ also, so that $\exp(\alpha(\hat{\nu})) = \pm 1$ for all α. We have noted in the proof of Lemma 9 that the conjugation C has the property $C\Lambda_\alpha = \Lambda_{-\alpha}$ for each root space Λ_α, and thus, if $\lambda_\alpha \in \Lambda_\alpha$, we see that $C\exp(\mathrm{Ad}(\hat{\nu}))\,C\lambda_\alpha = \exp(\mathrm{Ad}(\hat{\nu}))\,\lambda_\alpha$. Therefore $C\exp(\mathrm{Ad}(\hat{\nu}))\,C = \exp(\mathrm{Ad}(\hat{\nu}))$. We have already seen that $\Phi_C = \Phi$; thus $\hat{\tau}_C = \hat{\tau}$. It follows by Corollary 18, and this brings us close to our goal, that the conjugation τC is equivalent to the conjugation $\hat{\tau}C$.

We have $\alpha_1(\hat{\nu}) = \alpha_1(\Phi\hat{\nu}) = \alpha_5(\hat{\nu})$ and $\alpha_2(\hat{\nu}) = \alpha_2(\Phi\hat{\nu}) = \alpha_4(\hat{\nu})$; thus $\exp(\alpha_1(\hat{\nu})) = \exp(\alpha_5(\hat{\nu})) = \pm 1$ and $\exp(\alpha_2(\hat{\nu})) = \exp(\alpha_4(\hat{\nu})) = \pm 1$. Suppose that either $\exp(\alpha_3(\hat{\nu})) = -1$ or $\exp(\alpha_5(\hat{\nu})) = -1$; to be definite, let $\exp(\alpha_3(\hat{\nu})) = -1$. We saw (cf. the remark following Lemma II.5.9) that the automorphism ϕ_{α_3} defined by

$$\phi_{\alpha_3} = \exp(i\pi(2\alpha_3(\mu_{\alpha_3}))^{-1/2}\,\mathrm{Ad}(\lambda_{\alpha_3} + \lambda_{-\alpha_3}))$$

maps $\hat{M}^{(1)}$ into $\hat{M}^{(1)}$, and that $\phi_\alpha(\mu_\beta) = \mu_{W_{\alpha_3}\beta}$, where W_{α_3} is the Weyl reflection of Definition II.5.8. Since $C\lambda_{\pm\alpha_3} = -\lambda_{\pm\alpha_3}$, we have $C\phi_{\alpha_3} = \phi_{\alpha_3}C$. Since $\Phi\lambda_{\pm\alpha_3} = \lambda_{\pm\alpha_3}$, we have also $\Phi\phi_{\alpha_3} = \phi_{\alpha_3}\Phi$. Thus

$$\phi_{\alpha_3}\exp(\mathrm{Ad}(\hat{\nu}))\,\Phi C\phi_{\alpha_3}^{-1} = \phi_{\alpha_3}^2\exp(\mathrm{Ad}(\hat{\nu}))\,\Phi C.$$

If β is a root such that $\beta + \alpha$ is a root but neither $\beta - \alpha$ nor $\beta + 2\alpha$ is a root, then, by the Remark made in the semi-final paragraph under heading A' above, we find that the restriction $\phi_\alpha^2|\Lambda_\beta$ of ϕ_α^2 to the root space Λ_β is $-I$; on the other hand, if β is a root which is not a neighbor of α, we have $\phi_\alpha^2|\Lambda_\beta = I$. Moreover, we deduce from the formula for $\exp(it(\lambda_\alpha + \lambda_{-\alpha}))$ given in the course of the proof of Lemma II.5.9 that $\phi_\alpha^2|\Lambda_\alpha = I$. Since

$\phi_\alpha^2|\hat{M}^{(1)} = I$, we have $\phi_\alpha^2 = \exp(\mathrm{Ad}(\nu_3))$ where $\hat{\nu}_3 \in \hat{M}^{(1)}$ by Lemma II.5.11; and then, by what we have just said, we must have $\exp(\alpha(\nu_3)) = \pm 1$, the plus sign applying for $\alpha = \alpha_1, \alpha_3, \alpha_5$, and the minus sign applying for $\alpha = \alpha_2, \alpha_4, \alpha'$. Therefore, replacing $\hat{\tau}C$ by an equivalent conjugation if necessary, we may suppose without loss of generality that if $\exp(\alpha_3(\hat{\nu})) = -1$, then $\exp(\alpha'(\hat{\nu})) = -1$ also. Arguing similarly in case $\exp(\alpha'(\hat{\nu})) = -1$, we see that if $\exp(\alpha'(\hat{\nu})) = -1$, then we may suppose without loss of generality that $\exp(\alpha_3(\hat{\nu})) = -1$ also. Therefore, in any case, we may suppose without loss of generality that $\hat{\tau} = \exp(\mathrm{Ad}(\hat{\nu}))\,\Phi$, where $\exp(\alpha_1(\hat{\nu})) = \exp(\alpha_5(\hat{\nu})) = \pm 1$, $\exp(\alpha_2(\hat{\nu})) = \exp(\alpha_4(\hat{\nu})) = \pm 1$, and $\exp(\alpha_3(\hat{\nu})) = \exp(\alpha'(\hat{\nu})) = \pm 1$.

Yet another reduction can be made. Suppose, for the sake of definiteness, that $\exp(\alpha_1(\hat{\nu})) = -1$. Choose $\nu_2 \in \hat{M}^{(1)}$ so that $\alpha_5(\nu_2) = \pi i$ and $\alpha_1(\nu_2) = \alpha_2(\nu_2) = \alpha_3(\nu_2) = \alpha_4(\nu_2) = \alpha'(\nu_2) = 0$. Then since $C\Lambda_\alpha = \Lambda_{-\alpha}$ for every root α, we have $C \exp(\mathrm{Ad}(\nu_2))\, C\lambda_\alpha = \lambda_\alpha$ for $\alpha = \pm\alpha_i$, $i = 1, \ldots, 5$ and also for $\alpha = \pm\alpha'$, so that $C \exp(\mathrm{Ad}(\nu_2))\, C = \exp(\mathrm{Ad}(\nu_2))$. Therefore $\hat{\tau}C$ is equivalent to $\exp(\mathrm{Ad}(\hat{\nu} + \nu_2 - \Phi\nu_2))\,\Phi C$. But $\exp(\alpha(\hat{\nu} + \nu_2 - \Phi\nu_2))\,\Phi C = \pm\exp(\alpha(\hat{\nu}))$, the plus sign to be applied if $\alpha = \alpha_2, \alpha_3, \alpha_4$, or α', while the minus sign is to be applied if $\alpha = \alpha_1$ or $\alpha = \alpha_5$. Arguing in the same way if $\exp(\alpha_2(\hat{\mu})) = -1$, we see that, replacing $\hat{\tau}C$ by an equivalent conjugation if necessary, we may suppose without loss of generality that $\hat{\tau} = \exp(\mathrm{Ad}(\hat{\nu}))\,\Phi$, where $\exp(\alpha_i(\hat{\nu})) = 1$, $i = 1, 2, 4, 5$, while $\exp(\alpha_3(\hat{\nu})) = \exp(\alpha'(\hat{\nu})) = \pm 1$.

This shows that there are at most two non-isomorphic real subalgebras of $\hat{\Lambda}^{(1)}$ defined by conjugations of the form τC, where the automorphism τ is outer. One of these two must of course be the Lie algebra RE_6 of signature 42 which was noted above. If we let C_0 be the conjugation of $\hat{\Lambda}^{(1)}$ defined by $\lambda \to \bar{\lambda}$, $v \to \bar{v}$, $v' \to \bar{v}'$, let λ_0 be the 9×9 matrix which may be partitioned as

$$\begin{pmatrix} -\pi i & & & \\ & -\pi i & & \\ & & i(I - \mathcal{Y}_6) & \\ & & & -\pi i \end{pmatrix}$$

(cf. formula (1) above, under heading A following Lemma 20), so that $\exp(\lambda_0)$ is partitioned as

$$\begin{pmatrix} -1 & & & \\ & -1 & & \\ & & \mathcal{Y}_6 & \\ & & & -1 \end{pmatrix},$$

and let QE_6 be the real Lie algebra defined by

$$QE_6 = \{\hat{\lambda} \in \hat{\Lambda}^{(1)} | \exp(\mathrm{Ad}(\lambda_0))\, C_0 \hat{\lambda} = \hat{\lambda}\},$$

then we verify easily, using the description of the natural bilinear form of $\hat{\Lambda}^{(1)}$ given above in the first paragraph under the present heading C' that the signature of QE_6 is 36.

Our analysis of the conjugations in E_6 is therefore complete.

E′. Analysis of the conjugations in the Lie algebra of type F_4

We saw in Section II.4 that if $\hat{\Lambda}$ is a Lie algebra of type F_4 and $\hat{M}$ its four-dimensional Cartan subalgebra, then the non-zero roots of $\hat{M}$ in $\hat{\Lambda}$ have the form $\pm\omega_i \pm \omega_j$, where $i \neq j$ and $1 \leqslant i, j \leqslant 4$; $\pm\omega_i$, where $1 \leqslant i \leqslant 4$, and $\frac{1}{2}(\pm\omega_1 \pm \omega_2 \pm \omega_3 \pm \omega_4)$. We saw in the course of the proof of Lemma 9 that the positive conjugation C of that lemma could be chosen so that $C\hat{M} = \hat{M}$, and such that $C\Lambda_\alpha = \Lambda_{-\alpha}$ for each root α of $\hat{M}$ in $\hat{\Lambda}$. The Dynkin diagram of type F_4 admits no symmetries, and thus, by Lemma 12, every automorphism of $\hat{\Lambda}$ is inner. The set Π consisting of the roots $\omega_1 - \omega_2, \omega_2 - \omega_3, \omega_3, -\frac{1}{2}(\omega_1 + \omega_2 + \omega_3 + \omega_4)$ is a simple set of roots. It will be convenient for us in what follows to call these roots $\alpha_1, \alpha_2, \alpha_3, \alpha_4$.

By Corollary 16, every conjugation C_1 in $\hat{\Lambda}$ is equivalent to a conjugation of the form ϕC, where ϕ is an automorphism, $\phi^2 = \phi$, and $\phi_C = \phi$. By Lemma 32 and Corollary 35, there exists an automorphism of the form $\psi = \exp(\mathrm{Ad}(\nu))$, where $\nu \in \hat{M}$, and an automorphism τ, such that $\tau\phi\tau^{-1} = \exp(\mathrm{Ad}(\nu))$. Thus $\psi^2 = I$, and it follows that $\exp(\alpha(\nu))\, \lambda_\alpha = \psi\lambda_\alpha = \pm\lambda_\alpha$ for each root α of $\hat{M}$ in $\hat{\Lambda}$ and each λ_α in the corresponding root space Λ_α. Since $C\Lambda_\alpha = \Lambda_{-\alpha}$ for each root α, we find easily that $C\psi C = \psi$, and it follows by Corollary 18 that the conjugations ϕC and ψC are equivalent. By Corollary II.3.19 (*iv*), and since $\mathrm{Ad}(\nu)\, \lambda_\alpha = \alpha(\nu)\, \lambda_\alpha$ if λ_α belongs to the root space Λ_α, the automorphism ψ is determined by the four numbers $\exp(\alpha_i(\nu))$, each of which is ± 1; thus ψ must be one of sixteen automorphisms, each of which is in correspondence with one of the sixteen possible symbolic quadruples $(\pm 1, \pm 1, \pm 1, \pm 1)$ of $+1$'s and -1's.

If α is a root such that $\psi\lambda_\alpha = -\lambda_\alpha$ for λ_α in the root space Λ_α, and if we define ϕ_α as in the Remark made above in the semifinal paragraph under heading A', we find that we have $\phi_\alpha\psi C\phi_\alpha^{-1} = \phi_\alpha\psi\phi_\alpha^{-1} C = \phi_\alpha^2\psi C$. We noted in the cited Remark that $\phi_\alpha^2 = \exp(\mathrm{Ad}(\eta))$ with $\eta \in \hat{M}$; thus $\phi_\alpha^2\psi = \exp(\mathrm{Ad}(\eta + \nu))$. Since $(\phi_\alpha^2\psi)^2 = (\phi_\alpha\psi\phi_\alpha^{-1})^2 = I$, we must have $\exp(\beta(\eta + \nu)) = \pm 1$ for each non-zero root β. Therefore, the transformations $\psi C \to \phi_\alpha\psi C\phi_\alpha^{-1}$ map our family of sixteen conjugations into itself. By the Remark made above in the semi-final paragraph under heading A', we have $\exp(\beta(\eta)) = 1$ if $\beta = \alpha$ or if β is not a neighbor of α, while $\exp(\beta(\eta)) = -1$ if β is a simple neighbor of α; using this observation, we shall show that all sixteen of the above conjugations are equivalent to three among them.

Our computations will be as terse as possible if we introduce the notation σ_i for the quantity $\exp(\alpha_i(\nu))$, denote each conjugation by the corresponding symbol $(\sigma_1, \sigma_2, \sigma_3, \sigma_4)$, and indicate equivalence of conjugations by the writing $(\sigma_1, \sigma_2, \sigma_3, \sigma_4) \sim (\sigma_1', \sigma_2', \sigma_3', \sigma_4')$. We proceed as follows. Suppose that one of σ_1, σ_2 is -1. Then, by the preceding paragraph, we may reverse the other; consequently, we have only to consider conjugations for which $\sigma_1 = \sigma_2$. Similarly if $\sigma_4 = -1$, we may assume that $\sigma_3 = -1$ also. For the same reason $(-1, -1, -1, 1) \sim (-1, \pm 1, -1, -1) \sim (-1, -1, -1, -1)$, and $(1, 1, -1, 1)$ is equivalent either to $(1, 1, -1, -1)$, or to $(1, -1, -1, -1) \sim (-1, -1, 1, -1) \sim (-1, -1, -1, -1)$. Therefore, every conjugation is equivalent to one of the four conjugations for which $\sigma_1 = \sigma_2$ and $\sigma_3 = \sigma_4$. Finally, $(-1, -1, 1, 1) \sim (1, -1, -1, 1)$, which is in turn equivalent either to $(1, 1, -1, -1)$, or to $(1, -1, -1, 1) \sim (-1, -1, 1, -1) \sim (-1, -1, -1, -1)$. Therefore, as stated, only three of our sixteen conjugations are mutually inequivalent.

If we put $\psi C = C$, which corresponds to $(1, 1, 1, 1)$, we have, of course, a real Lie algebra $LF_4^{(0)}$, of signature 0. To compute the two other possible signatures, we argue as follows. We may choose $\lambda_{\pm\alpha}$ in the root spaces $\hat{\Lambda}_{\pm\alpha}$ in such a way that $(\lambda_\alpha, \lambda_{-\alpha}) = 1$ and $C\lambda_\alpha = -\lambda_{-\alpha}$. First letting ψ correspond to the symbol $(-1, -1, -1, -1)$, we see that $\psi\lambda_\beta = \lambda_\beta$ if β is the sum of an even number of elements of Π, while $\psi\lambda_\beta = -\lambda_\beta$ if β is the sum of an odd number of elements of Π. The signature of the corresponding real Lie algebra $LF_4^{(1)}$ is then twice the number of roots β of the latter kind, which number is easily verified to be 18, so that the signature is 36. In the same way, letting ψ be the automorphism corresponding to the symbol $(1, 1, -1, -1)$, we find that $\psi\lambda_\beta = -\lambda_\beta$ for exactly 24 non-zero roots, so that the signature of the corresponding real Lie algebra $LF_4^{(2)}$ is 24.

Our labors are now at an end: we may state the following theorem, which summarizes the classification of real simple Lie algebras.

36. Theorem. The real simple Lie algebras are as follows:

I. The real Lie algebras obtained from the various complex simple Lie algebras of Theorem II.3.36 by restricting the field of scalars from the complex field to the real field.

II. The various real simple Lie algebras of "classical" type listed in Theorem 21 of the present section.

III. The following real simple Lie algebras of "exceptional type" described in the preceding paragraphs:

(a) Five real Lie algebras $HE_6^{(6)}$, $HE_6^{(5)}$, $HE_6^{(4)}$, RE_6, and QE_6, of signatures 0, 32, 40, 42, and 36 respectively;

(b) Four real Lie algebras $HE_7^{(7)}$, $HE_7^{(6)}$, $HE_7^{(5)}$, and $HE_7^{(4)}$ of signatures 0, 54, 64, and 70 respectively;

(c) Three real Lie algebras $HE_8^{(9)}$, $HE_8^{(7)}$, and $HE_8^{(5)}$ of signatures 0, 112, and 128 respectively;

(d) Three Lie algebras $LF_4^{(0)}$, $LF_4^{(1)}$, and $LF_4^{(2)}$ of signatures 0, 36, and 24 respectively;

(e) Two real Lie algebras $HG_2^{(3)}$ and $HG_3^{(1)}$ of signatures 0 and 8 respectively.

None of these real Lie algebras are isomorphic.

The representation theory of real simple algebras offers no additional difficulties. To describe the essential results, we require the following terminology, however.

37. Definition. (a) Let Λ be a real Lie algebra, and let ϱ_1 be a representation of Λ in complex vector space V_1.

Let $v_1, \ldots, v_n$ be a basis for V_1; then the *complex conjugate representation* $\bar{\varrho}_1$ of the representation ϱ_1 is defined by the condition that the matrix $(\bar{\varrho}_1)_{ij}(\lambda)$ of $\bar{\varrho}_1(\lambda)$ in the basis $v_1, \cdots, v_n$ is the complex conjugate of the matrix of $\varrho_1(\lambda)$ in this same basis.

(b) Let Λ_1 be a Lie algebra, and let ϱ_1 and ϱ_2 be representations of Λ_1 in the complex vector spaces V_1 and V_2 respectively. Let V_1' and V_2' be the spaces of all linear functionals on V_1 and V_2 respectively, and let $V_1 \otimes V_2$ be the set of all functions $\Phi(v_1', v_2')$, $v_1' \in V_1'$, $v_2' \in V_2'$, which are linear in both variables v_1' and v_2'. Put

$$(\varrho_1'(\lambda)\, v_1')\,(v_1) = -v_1'(\varrho(\lambda)\, v_1), \quad v_1 \in V_1, \quad v_1' \in V_1', \quad \lambda \in \Lambda_1;$$

$$(\varrho_2'(\lambda)\, v_2')\,(v_2) = -v_2'(\varrho(\lambda)\, v_2), \quad v_2 \in V_2, \quad v_2' \in V_2', \quad \lambda \in \Lambda_1;$$

and

$$\{((\varrho_1 \otimes \varrho_2)(\lambda)\Phi\}(v_1', v_2') = -\Phi(\varrho_1'(\lambda)\, v_1', v_2') - \Phi(v_1', \varrho_2'(\lambda)\, v_2'),$$

$$v_1' \in V_1', \quad v_2' \in V_2', \quad \Phi \in V_1 \otimes V_2, \quad \lambda \in \Lambda_1.$$

The representation $\varrho_1 \otimes \varrho_2$ of Λ_1 thus defined is called the *Kronecker product* representation of the two representations ϱ_1 and ϱ_2.

Concerning (a) of Definition 37, we leave it to the reader to verify that if another basis than $v_1, \ldots, v_n$ is used to define the complex conjugate representation, we obtain a representation of Λ which is equivalent to $\bar{\varrho}_1$.

38. Theorem. (a) Let Λ be the real simple Lie algebra

$$\Lambda = \{\tilde{\lambda} \in \hat{\Lambda} | C\,\tilde{\lambda} = \tilde{\lambda}\},$$

where $\hat{\Lambda}$ is a complex simple Lie algebra and C is a conjugation in $\hat{\Lambda}$. Then the irreducible representations of Λ are the restrictions to Λ of the irreducible representations of $\hat{\Lambda}$; two such restrictions are equivalent if and only if the representations of $\hat{\Lambda}$ from which they are derived are equivalent.

(b) Let Λ_1 be a complex simple Lie algebra, and let Λ be the real simple Lie algebra derived from Λ_1 by restricting the field of scalars from the complex field to the real field. Then the irreducible representations of Λ are the representations of the form $\varrho_1 \otimes \bar{\varrho}_2$, where ϱ_1 and ϱ_2 are representations of Λ which are at the same time representations of Λ_1. Two such representations $\varrho_1 \otimes \bar{\varrho}_2$ and $\varrho_3 \otimes \bar{\varrho}_4$ are equivalent if and only..if ϱ_1 is equivalent to $\bar{\varrho}_3$ and ϱ_2 is equivalent to $\bar{\varrho}_4$.

Proof: (a) Let ϱ be an irreducible representation of Λ in a space V. We may write each $\hat{\lambda} \in \Lambda$ uniquely as $\hat{\lambda} = \frac{1}{2}(\hat{\lambda} + C\hat{\lambda}) + \frac{1}{2}(\hat{\lambda} - C\hat{\lambda})$; it is easily verified that $\hat{\varrho}(\hat{\lambda}) = \varrho\left(\dfrac{1}{2}(\hat{\lambda} + C\tilde{\lambda})\right) + i\varrho\left(\dfrac{1}{2i}(\hat{\lambda} - C\hat{\lambda})\right)$ is an extension of ϱ to all of $\hat{\Lambda}$. Plainly, $\hat{\varrho}$ is irreducible. Suppose conversely that ϱ is the restriction to Λ of an irreducible representation $\hat{\varrho}$ of $\hat{\Lambda}$. If ϱ is not irreducible, there exists a non-zero vector v in V, and a non-zero linear functional v' on V, such that $v'(\varrho(\lambda)\, v) = 0$ for all $\lambda \in \Lambda$. Then $v'(\hat{\varrho}(z\hat{\lambda} + Cz\hat{\lambda})\, v) = 0$ for $\hat{\lambda} \in \hat{\Lambda}$ whenever z is real; and hence for all z. Thus, putting $z = i$ and $z = 1$ and adding we find $v'(\hat{\varrho}(\hat{\lambda})\, v) = 0$, $\hat{\lambda} \in \hat{\Lambda}$, contradicting the irreducibility of $\hat{\varrho}$. This shows that the irreducible representations of Λ are the restrictions to Λ of the irreducible representations of $\hat{\Lambda}$. Plainly, if two representations of $\hat{\Lambda}$ are equivalent, their restrictions to Λ are equivalent also. Conversely, let ϱ_1 and ϱ_2 be representations of $\hat{\Lambda}$ in vector spaces V_1 and V_2, and suppose that their restrictions to Λ are equivalent. Then there exists a one-to-one linear mapping T of V_1 onto V_2 such that $\varrho_1(\lambda) = T^{-1}\varrho_2(\lambda)\,T$ for each $\lambda \in \Lambda$. But then evidently $\varrho_1(i\lambda) = T^{-1}\varrho_2(i\lambda)\,T$ for $\lambda \in \Lambda$, and since each $\nu \in \hat{\Lambda}$ can be written as $\nu = \lambda + i\mu$, where λ and μ lie in Λ, ϱ_1 and ϱ_2 are equivalent, completing the proof of (a).

(b) Note that, in proving part (a) of the present lemma, we did not use the simplicity of Λ at all, but merely the fact that C was a conjugation in $\hat{\Lambda}$, and that Λ was $\{\hat{\lambda} \in \hat{\Lambda} | C\hat{\lambda} = \hat{\lambda}\}$. If Λ_1 is a complex simple Lie algebra, and Λ the real simple Lie algebra derived from it by restricting the field of scalars, we may let $\hat{\Lambda}$ be the direct sum $\hat{\Lambda} = \Lambda_1 \oplus \Lambda_1$ of two isomorphic copies of Λ_1, and, using Lemma 9, let C_1 be a conjugation in Λ_1. Then $\hat{\Lambda}$ is the set of all pairs $\{\lambda, \mu\}$ with $\lambda, \mu \in \Lambda_1$, and we can define a conjugation C in $\hat{\Lambda}$ by $C\{\lambda, \mu\} = \{C_1\mu, C_1\lambda\}$. The real Lie algebra $\{\hat{\lambda} \in \hat{\Lambda} | C\hat{\lambda} = \hat{\lambda})$ is then the set of all pairs $\{\lambda, C\lambda\}$ with $\lambda \in \Lambda_1$; and the correspondence η: $\lambda \leftrightarrow \{\lambda, C\lambda\}$ is plainly an isomorphism between this real Lie algebra and the real Lie algebra Λ. It follows by the arguments used to prove (a) that the irreducible representations of Λ are the representations $(\varrho|\Lambda)\,\eta$, where ϱ denotes a generic irreducible representation of $\hat{\Lambda}$, and two such representations $(\varrho|\Lambda)\,\eta$ and $(\tilde{\varrho}|\Lambda)\,\eta$ are equivalent if and only if ϱ and $\tilde{\varrho}$ are equivalent.

It is plain, using Theorem II.2.3, that $\hat{\Lambda}$ is semi-simple. Let M_1 be a Cartan subalgebra of Λ_1; then the reader will verify easily, using Definition II.2.7, that the direct sum $M = M_1 \oplus M_1$ of two isomorphic copies of M_1 is a Cartan subalgebra of $\hat{\Lambda}$, and that the root spaces of $\hat{M}$ in $\hat{\Lambda}$ are in the first place $\hat{M}$, in the second place the direct sums $\Lambda_\alpha \oplus \{0\}$ and $\{0\} \oplus \Lambda_\alpha$, where α denotes a generic non-zero root of M_1 in Λ_1. If for each non-zero root α of M_1 in Λ_1 we let α' denote the linear functional on $\hat{M}$ defined by $\alpha'\{\lambda, \mu\} = \alpha(\lambda)$, and let α'' denote the linear functional on $\hat{M}$ defined by $\alpha''\{\lambda, \mu\} = \alpha(\mu)$, it follows that the set of roots of $\hat{M}$ in $\hat{\Lambda}$ consists of zero and of all α' and α''. It is plain from Definition II.2.2 that the natural bilinear form of $\hat{\Lambda}$ is expressed by $(\{\lambda, \mu\}, \{\lambda_1, \mu_1\}) = (\lambda, \lambda_1) + (\mu, \mu_1)$ in terms of the natural bilinear form of Λ_1. Thus we have $\mu_{\hat{\alpha}'} = \{\mu_\alpha, 0\}$ and $\mu_{\alpha''} = \{0, \mu_\alpha\}$ in the sense of Corollary II.3.8. If Π_1 is a simple set of roots of M_1 in Λ_1, and if we let $\hat{\Pi} = \{\alpha', \alpha'' | \alpha \in \Pi_1\}$, it then follows from Definition 3.17 that $\hat{\Pi}$ is a simple set of roots of $\hat{M}$ in $\hat{\Lambda}$.

Let ϱ_1 and ϱ_2 be irreducible representations of Λ_1 in spaces V_1 and V_2 respectively. Let V_1', V_2', $V_1 \otimes V_2$, ϱ_1', and ϱ_2' be as in part (b) of Definition 37, and let $\varrho_1 \circledast \varrho_2$ be the representation of $\hat{\Lambda}$ in $V_1 \otimes V_2$ defined by

$$[((\varrho_1 \circledast \varrho_2)(\{\lambda, \mu\}))\,\Phi](v_1', v_2') = -\Phi(\varrho_1'(\lambda)\,v_1', v_2') - \Phi(v_1', \varrho_2'(\mu)\,v_2'),$$
$$v_1' \in V_1', \quad v_2' \in V_2', \quad \Phi \in V_1 \otimes V_2, \quad \lambda, \mu \in \Lambda_1.$$

We shall show that $\varrho = \varrho_1 \circledcirc \varrho_2$ is irreducible. We proceed by contradiction. Suppose that ϱ is reducible. Then, by Theorem II.6.2 we may write $V_1 \otimes V_2$ as a direct sum $V_1 \otimes V_2 = X \oplus Y$, where both X and Y are invariant under ϱ. Define the linear transformation E in $V_1 \otimes V_2$ by requiring $E|X = I$, $E|Y = 0$. Then it is easily seen that $E\varrho(\hat{\lambda}) = \varrho(\hat{\lambda})\, E$ for each $\hat{\lambda} \in \hat{\Lambda}$. For each $v_1 \in V_1$ and $v_2 \in V_2$, and let $v_1 \otimes v_2$ denote the element of $V_1 \otimes V_2$ defined by

$$(v_1 \otimes v_2)\,(v_1', v_2') = v_1'(v_1)\, v_2'(v_2).$$

Then the reader will easily verify that

$$\varrho(\{\lambda, 0\})\,(v_1 \otimes v_2) = (\varrho_1(\lambda)\, v_1) \otimes v_2$$

and

$$\varrho(\{0, \lambda\})\,(v_1 \otimes v_2) = v_1 \otimes \varrho_2(\lambda)\, v_2$$

for each $\lambda \in \Lambda_1$. Choose fixed elements $v_2 \in V_2$ and $v_2' \in V_2'$; then the expression $(E(v_1 \otimes v_2))\,(v_1', v_2')$ is linear in V_1', so that there exists a unique $\tilde{v}_1 \in V_1$ such that $(E(v_1 \otimes v_2))\,(v_1', v_2') = v_1'(\tilde{v}_1)$. The correspondence $v_1 \to \tilde{v}_1$ is plainly unique, and thus defines a linear mapping F in V_1, so that we may write $(E(v_1 \otimes v_2))\,(v_1', v_2') = v_1'(Fv_1)$. Then

$$\begin{aligned}
v_1'(F\varrho_1(\lambda)\, v_1) &= (E(\varrho_1(\lambda)\, v_1 \otimes v_2))\,(v_1', v_2')\\
&= (E\varrho(\{\lambda, 0\})\,(v_1 \otimes v_2))\,(v_1', v_2')\\
&= (\varrho(\{\lambda, 0\})\, E(v_1 \otimes v_2))\,(v_1', v_2')\\
&= -(E(v_1 \otimes v_2))\,(\varrho_1'(\lambda)\, v_1', v_2') = -(\varrho_1'(\lambda)\, v_1')\,(Fv_1)\\
&= v_1'(\varrho_1)\,(\lambda)\,(Fv_1),
\end{aligned}$$

proving that $\varrho_1(\lambda)\, F = F\varrho_1(\lambda)$ for each $\lambda \in \Lambda_1$. But then each eigenspace of F is invariant for the representation ϱ_1, and, since ϱ_1 is an irreducible representation, it follows that F is a constant multiple of the identity. We may therefore write

$$(E(v_1 \otimes v_2))\,(v_1', v_2') = a(v_2', v_2)\, v_1'(v_1).$$

But then the constant a is seen to depend linearly on v_2' and v_2, so that arguing as above we see that there is a linear mapping G in V_2 such that

$$(E(v_1 \otimes v_2))\,(v_1', v_2') = v_2'(Gv_2)\, v_1'(v_1);$$

and continuing our argument as above, we find that there exists a constant b such that

$$(E(v_1 \otimes v_2))\,(v_1', v_2') = b v_2'(v_2)\, v_1'(v_1).$$

But then $E(v_1 \otimes v_2) = b v_1 \otimes v_2$. The reader will verify without difficulty that the elements of the form $v_1 \otimes v_2$ span $V_1 \otimes V_2$; therefore $E = bI$, contradicting one of the two statements $E|X = I$ and $E|Y = 0$. This contradiction proves our assertion that $\varrho_1 \textcircled{\otimes} \varrho_2$ is irreducible.

Let $v_1 \in V_1$ be a minimal weight vector (cf. Definition II.6.8) of ϱ_1 belonging to the minimal weight γ_1, and $v_2 \in V_2$ be a minimal weight vector of ϱ_2 belonging to the minimal weight γ_2. Define $\gamma(\{\lambda, \mu\}) = \gamma_1(\lambda) + \gamma_2(\mu)$ for $\{\lambda, \mu\} = \hat{M} = M_1 + M_1$. Then it is readily verified that $v_1 \otimes v_2$ is a minimal weight vector of $\varrho_1 \textcircled{\otimes} \varrho_2$, belonging to the minimal weight γ. Thus, using Theorem II.6.13 and the description of the simple set $\hat{\Pi}$ of roots of $\hat{M}$ in $\hat{\Lambda}$ given above, it follows that *every* irreducible representation of $\hat{\Lambda}$ is equivalent to a representation of the form $\varrho_1 \textcircled{\otimes} \varrho_2$, where ϱ_1 and ϱ_2 are irreducible representations of Λ_1, and that two such representations $\varrho_1 \textcircled{\otimes} \varrho_2$ and $\tilde{\varrho}_1 \textcircled{\otimes} \tilde{\varrho}_2$ are equivalent if and only if ϱ_1 is equivalent to $\tilde{\varrho}_1$ and ϱ_2 is equivalent to $\tilde{\varrho}_2$. It follows that every irreducible representation of Λ is equivalent to a representation of the form

$$\lambda \to \varrho_1(\lambda) \textcircled{\otimes} \varrho_2(C\lambda),$$

where ϱ_1 and ϱ_2 are irreducible representations of Λ_1, and even that two such representations

$$\lambda \to \varrho_1(\lambda)\ \varrho_2(C\lambda) \quad \text{and} \quad \lambda \to \tilde{\varrho}_1(\lambda)\ \tilde{\varrho}_2(C\lambda)$$

are equivalent if and only if ϱ_1 is equivalent to $\tilde{\varrho}_1$ and ϱ_2 is equivalent to $\tilde{\varrho}_2$.

This almost amounts to a proof of part (b) of the present theorem. To complete the proof of (b), we argue as follows. Let $v_1, \ldots, v_n$ be a basis for V_2, and, for each representation ϱ_2 of Λ_1 in V_2, let ϱ_2^+ denote the representation of Λ_1 in V_2 defined by the condition that the matrix $(\varrho_2^+(\lambda))_{ij}$ of $\varrho^+(\lambda)$ in the basis $v_1, \ldots, v_n$ is the complex conjugate $(\varrho_2'(C'\lambda'))_{ij}$ of the matrix of $\varrho_2(C\lambda)$ in this same basis. Then plainly $(\varrho_2^+)^+ = \varrho_2$, and it is easily seen that ϱ_2 is irreducible if and only if ϱ_2^+ is irreducible. The representation $\lambda \to \varrho_1(\lambda) \textcircled{\otimes} \varrho_2(C\lambda)$ of Λ is, by Definition 37, identical with the representation $\varrho_1 \otimes \varrho_2'^{+}$ of Λ. Thus we learn that if ϱ_1 and ϱ_2 are irreducible representations of $\hat{\Lambda}$, then $\varrho_1 \otimes \varrho_2'^{+}$ is an irreducible representation of Λ; that every irreducible representation of Λ is equivalent to a representation having this special form; and that if $\varrho_1 \otimes \varrho_2'^{+}$ is equivalent to $\tilde{\varrho}_1 \otimes \tilde{\varrho}_2'^{+}$, then ϱ_1 is equivalent to $\tilde{\varrho}_1$ and ϱ_2 is equivalent to $\tilde{\varrho}_2$. Replacing ϱ_2^+ by ϱ_2 and using $(\varrho_2^+)^+ = \varrho_2$, statement (b) of our theorem follows. Q.E.D.

2. Compact Connected Lie Groups

As a dividend of our drudgery in section 1, we obtain in the present section a theorem describing the structure of an arbitrary compact connected Lie group.

1. Lemma. Let G be a connected compact Lie group, $\Lambda(G)$ its Lie algebra. Then $\Lambda(G)$ is the direct sum $\Lambda(G) = \Lambda_1 \oplus \cdots \oplus \Lambda_n$, each of which is either a one-dimensional Lie algebra or a real simple Lie algebra of one of the types $HA_n^{(0)}$, $n \geqslant 1$;

$$LB_m^{(0)}, m \geqslant 2; HC_m^{(0)}, m \geqslant 3; LD_m^{(0)}, m \geqslant 4; HE_6^{(6)}; HE_7^{(7)}; HE_8^{(9)};$$
$$LF_4^{(0)}; \text{and } HG_2^{(3)}.$$

Proof. Let $\varrho(g) : \lambda \to g\lambda g^{-1}$ be the natural representation of G as a group of linear transformations of its Lie algebra. Then $\varrho(g)$ depends continuously on g by Lemma I.3.8(10). Introduce any arbitrary positive definite quadratic form $\{\lambda, \mu\}_0$ into $\Lambda(G)$ (say, by using a basis). Then put

$$\{\lambda, \mu\} = \int_G \{\varrho(g)\,\lambda, \varrho(g)\,\mu\}_0 \, dg, \quad \lambda, \mu \in \Lambda(G),$$

where the integral is a Haar integral. From the invariance of the Haar integral (cf. Lemma 0.65) we find that $\{\varrho(g)\,\lambda, \varrho(g)\,\mu\} = \{\lambda, \mu\}$, $\lambda, \mu \in \Lambda(G)$. Plainly, $\{\lambda, \mu\}$ is a positive definite quadratic form on $\Lambda(G)$. Differentiating the equation $\{\varrho(\exp(tv))\,\lambda, \varrho(\exp(tv)\,\mu\} = \{\lambda, \mu\}$ and putting $t = 0$, we find using Lemma I.5.3 that $\{\mathrm{Ad}(v)\,\lambda, \mu\} = -\{\lambda, \mathrm{Ad}(v)\,\mu\}$, $\lambda, \mu, v \in \Lambda(G)$. Thus, if Λ' is an ideal of $\Lambda(G)$, so is $\Lambda'' = \{\lambda \in \Lambda(G) | \{\lambda, \lambda'\} = 0$ for all $\lambda \in \Lambda'\}$; and $\Lambda(G) = \Lambda' \oplus \Lambda''$. An evident dimensional induction now shows that $\Lambda(G)$ is a direct sum $\Lambda(G) = \Lambda_1 \oplus \cdots \oplus \Lambda_n$ of ideals satisfying $\{\Lambda_i, \Lambda_j\} = 0$ for $i \neq j$, each of which, regarded as a Lie algebra in its own right, is either one-dimensional or simple. If Λ_j is simple, then, arguing exactly as in the proof of Lemma I.10, we see that its natural bilinear form (λ, μ) is a constant multiple of the restriction to Λ_j of the form $\{\lambda, \mu\}$. Thus the signature of Λ_j is either zero or egual to the dimension of Λ_j. But, checking through the list of real algebras given by Theorem I.36 of the preceding section (cf. Theorem II.4.4 for the dimensions) we see that the signature of Λ_j is never equal to the dimension of Λ_j, and is equal to zero only if Λ_j is of one of the types listed in the statement of the present Lemma. Q.E.D.

If the signature of a real simple Lie algebra Λ is zero, then the connected component A_0 of the identity in the group A of all automorphisms of Λ, which by Lemma I.2.3 is a Lie group whose Lie algebra is Λ, is clearly a closed subgroup of A. But A is a closed subgroup of the group of all linear transformations in Λ which preserve the positive definite bilinear form $-(\lambda, \mu)$. Thus A is compact, and therefore A_0 is a compact Lie group. The following lemma shows that the simply connected covering group of A_0 is compact also, so that the unique simply connected Lie group defined by a real simple Lie algebra of signature zero is compact.

2. Lemma. Let G be a compact connected Lie group whose Lie algebra is simple. Then the simply connected covering group G_1 of G is also compact.

Proof. Let p be the natural homomorphism of G_1 onto G (cf. Theorem I.2.10), so that by this theorem the subgroup $H = p^{-1}(e)$ of G_1 is abelian and consists exclusively of isolated points, and $hg = gh$ for each $h \in H$ and $g \in G_1$. We shall show that H is finite. We proceed as follows. By the results of Section 0, there exists for each $g \in G$ an open set $U_g \subseteq G_1$ mapped homeomorphically by p onto a neighborhood of g, and thus a compact subset K_g of G_1 with interior $\dot{K}_g$ such that $p(\dot{K}_g)$ includes a neighborhood of g. Since G is compact, it therefore follows that there exists a compact subset K of G_1 with interior $\dot{K}$ such that $p(\dot{K}) = G$. Replacing K by $K \cup K^{-1}$ if necessary, we can suppose without loss of generality that $K = K^{-1}$.

Since $p(K) = G$ and $H = p^{-1}(e)$, it follows that $\dot{K}H = G_1$; and thus a finite number $\dot{K}h_1, \ldots, \dot{K}h_n$ of the sets $\dot{K}h$ will cover the compact set KK^{-1}. Let H_1 be the subgroup of H generated by its elements $h_1, \ldots, h_n$. Let p_1 be the natural homomorphism of G_1 onto G_1/H_1, (cf. Theorem I.2.11). Then since $KK^{-1} \subseteq KH_1$, we have $p_1(K)p_1(K)^{-1} \subseteq p_1(K)$; thus $p_1(K)$ is a subgroup of G_1/H_1. By Theorem I.2.11, $p_1(K)$ is open; thus $p_1(K)$ is an open subgroup of G_1/H_1. But G_1, and hence $p_1(G_1) = G_1/H_1$, is connected; therefore, by Lemma I.2.2, $p_1(K) = G_1/H_1$. Since K is compact, it follows that G_1/H_1 is compact.

Let $h \in H$, and let U be a neighborhood of h containing no other point of H, chosen so small that p_1 maps U homeomorphically onto an open set in G_1/H_1. If $p_1(U)$ contains a point $p_1(h') \neq p_1(h)$ with $h' \in H$, then we have $p_1(u) = p_1(h')$ for some $u \in U$, so that $u \in h'H_1 \subseteq H$, while $p(u) \neq p(h)$ so $u \neq h$, a contradiction. This shows that $p_1(U) \cap p_1(H) = \{h\}$, so that the subgroup $p_1(H)$ of G_1/H_1 consists of isolated points exclusively.

But G_1/H_1 is compact, and thus $p_1(H)$ is finite. To show that H is finite, it is therefore sufficient to show that H_1 is finite. Since H_1 is abelian and generated by $h_1, \ldots, h_n$, it is even sufficient to show that each of these elements is of finite order.

Suppose then, to proceed by contradiction, that not every h_j is of finite order. Let N be the set of all n-tuples $[p_1, \ldots, p_n]$ such that $h_1^{p_1} \cdots h_n^{p_n} = e$, so that N is closed under addition and subtraction, and let $\tilde{N}$ be the set of all rational multiples of the n-tuples of N, so that $\tilde{N}$ is a rational linear space. If $\tilde{N}$ were n-dimensional, then for k sufficiently large every n-tuple of the form $[0, 0, \ldots, 0, k!, 0, \ldots, 0]$ would belong to N, and then we would have $h_j^{k!} = e$ for each j, contradicting our assumption. Therefore $\tilde{N}$ must be at most $n-1$ dimensional, so that there exists a non-zero rational n-tuple $[q_1, \ldots, q_n]$ such that $q_1p_1 + \cdots + q_np_n = 0$ if $[p_1, \ldots, p_n] \in N$. Put $\phi(h) = q_1p_1 + \cdots + q_np_n$ if $h = h_1^{p_1}h_2^{p_2} \cdots h_n^{p_n} \in H_1$; then $\phi(h)$ is a real valued, non-constant function defined on H_1 and satisfying the functional equation $\phi(hh') = \phi(h) + \phi(h')$, $h, h' \in H_1$. We shall now show, and this is the heart of the present proof, that ϕ can be extended to a continuous function defined on all G_1 and satisfying the same functional equation.

We reason as follows. Arguing exactly as in the first paragraph of the present proof, we find a compact set $K_1 \subseteq G_1$ such that $K_1 = K_1^{-1}$ and such that $p_1(K_1) = G_1/H_1$. We then find a second compact set K_2 whose interior $\dot{K}_2$ includes K_1, and a continuous real function f_0 on G_1 which is non-negative, identically 1 on K_1, and identically zero outside of $\dot{K}_2$. If $g_1 \in G_1$, we have $p(g_1) = p(k_1)$ for some $k_1 \in K_1$; this shows that $g_1 = k_1h_1$ for some $k_1 \in K_1$ and $h_1 \in H_1$. On the other hand, since K_2 is compact and H_1 contains only isolated points, $g_1^{-1}K_2 \cap H_1$ must be finite. Thus, for each g_1 only finitely many of the terms in the sum

$$\sigma_0(g_1) = \sum_{h \in H_1} f_0(g_1h^{-1})$$

are non-vanishing, while at least one is positive. We have therefore $\sigma_0(g_1) > 0$, $g_1 \in G_1$, σ_0 is continuous, and $\sigma_0(g_1h_1) = \sigma_0(g_1)$ $g_1 \in G_1$, $h_1 \in H_1$. If we put $f(g) = f_0(g)/\sigma_0(g)$, we have consequently

$$\sum_{h \in H_1} f(gh^{-1}) \equiv 1, \quad g \in G_1.$$

Now put

$$\psi_0(g) = \sum_{h \in H_1} f(gh^{-1})\,\phi(h) + \sum_{h \in H_1} f(h)\,\phi(h),$$

so that ψ_0 is a continuous function defined on all G_1. If $h_1 \in H_1$ we have

$$\begin{aligned}\psi_0(h_1) &= \sum_{h\in H_1} f(h_1h^{-1})\,\phi(h) + \sum_{h\in H_1} f(h)\,\phi(h)\\ &= \sum_{h\in H_1} f(h)\,\phi(h_1h^{-1}) + \sum_{h\in H_1} f(h)\,\phi(h)\\ &= \phi(h_1)\sum_{h\in H_1} f(h) - \sum_{h\in H_1} f(h)\,\phi(h) + \sum_{h\in H_1} f(h)\,\phi(h)\\ &= \phi(h_1).\end{aligned}$$

Thus ψ_0 extends ϕ from H_1 to all of G_1.

In the same way, if $g \in G_1$ and $h_1 \in H_1$ we have

$$\begin{aligned}\psi_0(gh_1) &= \sum_{h\in H_1} f(gh)\,\phi(h_1h^{-1}) + \sum_{h\in H_1} f(h)\,\phi(h)\\ &= \phi(h_1) + \sum_{h\in H_1} f(gh)\,\phi(h^{-1}) + \sum_{h\in H_1} f(h)\,\phi(h)\\ &= \phi(h_1) + \psi_0(g) = \psi_0(h_1) + \psi_0(g).\end{aligned}$$

Thus for each fixed $g_1 \in G_1$ the function $\psi_1(g_1; g) = \psi_0(g_1g) - \psi_0(g)$ satisfies

$$\begin{aligned}\psi_1(g_1; gh) &= \psi_0(g_1gh) - \psi_0(gh)\\ &= \psi_0(g_1g) - \psi_0(g) + \psi_0(h) - \psi_0(h)\\ &= \psi_1(g_1; g), \quad h \in H_1.\end{aligned}$$

There consequently exists a function $\psi_2(g_1; \gamma)$ on the compact group G_1/H_1 such that $\psi_2(g_1; p_1(g)) = \psi_1(g_1; g)$; the reader will verify easily, that the function ψ_2 is continuous on G_1/H_1, and in the "parameter" g_1 as well. Now put

$$\psi(g_1) = \int_{G_1/H_1} \psi_2(g_1; \gamma)\, d\gamma$$

(the integration being with respect to the Haar measure of the compact group G_1/H_1).

If $h_1 \in H_1$, we have $\psi_1(h_1; g) = \psi_0(h_1g) - \psi_0(g) = \phi(h_1)$, and thus $\psi(h_1) = \phi(h_1)$, so that $\psi(h_1)$ is a continuous extension of ϕ to the whole group G_1. If $g_1, g_2 \in G_1$, we have

$$\psi_1(g_1g_2; g) = \psi_0(g_1g_2g) - \psi_0(g) = \psi_1(g_1; g_2g) + \psi_1(g_2; g),$$

and thus

$$\psi_2(g_1g_2; p_1(g)) = \psi_1(g_1; p_1(g_2g)) + \psi_1(g_2; p_1(g)),$$

or

$$\psi_2(g_1g_2; \gamma) = \psi_2(g_1; p_1(g_2)\gamma) + \psi_1(g_2; \gamma), \quad \gamma \in G_1/H_1.$$

Integrating this last identity with respect to γ, and using the invariance of the Haar measure (cf. Lemma 0.65) we find that $\psi(g_1g_2) = \psi(g_1) + \psi(g_2)$.

Since ψ extends ϕ, ψ is not identically zero. By Theorem I.4.14, ψ is C^∞; and thus, by Theorem I.3.12, ψ determines a non-trivial Lie homomorphism ψ_* of the Lie algebra Λ of G_1 into the Lie algebra Λ_0 of the additive Lie group of reals. Since Λ is also the Lie algebra of G (cf. Theorem I.2.11), Λ is simple; and thus the ideal $\psi_*^{-1}(0)$ of Λ must be zero, so that ψ_* is an isomorphism. Since $[\lambda, \lambda'] = 0$ if $\lambda, \lambda' \in \Lambda_0$, this is impossible. We have now a contradiction to our earlier assumption that not every h_j is of finite order; and our lemma is completely proved. Q.E.D.

We may now state the following theorem, which, while it still leaves room for improvement, is reasonably decisive.

3. Theorem. Let G be a compact connected Lie group. Then G is isomorphic to a factor group G^*/H^*, where G^* is a compact connected Lie group locally isomorphic to G, and H^* is a finite subgroup of G^*, such that $h^*g^* = g^*h^*$ if $g^* \in G^*$ and $h^* \in H^*$. Moreover, G^* is a direct product of finitely many groups chosen from the following list (repetitions allowed): C, the multiplicative group of complex numbers of modulus 1; and UA_n, $n \geqslant 1$; UB_n, $n \geqslant 2$; UC_n, $n \geqslant 3$; UD_n, $n \geqslant 4$; UE_6; UE_7; UE_8; UF_4; and UG_2; where these last groups are the simply connected compact groups whose Lie algebras are

$$HA_m^{(0)},\ LB_m^{(0)},\ HC_m^{(0)},\ LD_m^{(0)},$$

$$HE_6^{(6)},\ HE_7^{(7)},\ HE_8^{(9)},\ LF_4^{(0)},\ \text{and } HG_2^{(3)} \text{ respectively,}$$

in the notation introduced in Theorems 1.21 and 1.36.

Proof. By Lemma 1, and using the fact that if two simply connected Lie groups have the same Lie algebra they are isomorphic (cf. Corollary I.5.7) we see that the simply connected covering group G_1 of G has the form of a Kronecker product $E^n \times G_0$, where E^n is the additive group of an n-dimensional Euclidean space and G_0 is a finite Cartesian product of groups UA_n, UB_n, etc. Moreover, G is the factor group $G = G_1/H$ of G_1 by a subgroup H consisting exclusively of isolated points h such that $g_1h = hg_1$ for each $g_1 \in G_1$.

Let us now consider an element $g_0 \in G_0$ of the form $g_0 = \exp(\lambda_0)$, where $\lambda_0 \in \Lambda_0$, Λ_0 being the Lie algebra of G_0. We have $\exp(\nu_0)\, g_0 \exp(-\nu_0) = \exp(\exp(\nu_0)\,\lambda_0 \exp(-\nu_0))$ for each $\nu_0 \in \Lambda_0$, and thus if $g_0 g = g g_0$ for each $g \in G_0$ we have $\exp(t\nu_0)\,\lambda_0 \exp(-t\nu_0) = \lambda_0$ for each $\nu_0 \in \Lambda_0$ and each real t; differentiating and using Definition I.3.7, we find that $[\nu_0, \lambda_0] = 0$ for each $\nu_0 \in \Lambda_0$. Since Λ_0 is the direct sum of simple Lie algebras, it follows that $\lambda_0 = 0$ and $g_0 = e$. Since (cf. Lemma I.4.11) every g_0 in a small neighborhood U of the element e of G_0 has the form $\exp(\lambda_0)$, it follows that U contains no element $g_0 \neq e$ such that $g g_0 = g_0 g$ for each $g \in G_0$.

If we now make use of the relationship $G_1 = E^n \times G_0$ to write each $h \in H \subseteq G_1$ as $h = [v(h), h_0(h)]$, where $v(h) \in E^n$ and $h_0 \in G_0$, it follows that the range of the homomorphism $h \to h_0(h)$ must be a subgroup of G_0 containing only isolated points. Since G_0 is a Cartesian product of groups to which Lemma 2 applies, G_0 is compact, and thus the range of $h \to h_0(h)$ must be finite. Thus, if H_1 is the subgroup $\{h \in H \mid h_0(h) = e\}$ of H, the factor group H/H_1 is finite. We have $H_1 = \{[v(H_1), e]\}$, and if we call the subgroup $v(H_1)$ of E^n by the name V_1 and put $T^n = E^n/V_1$, then the reader will easily verify that G_1/H_1 is isomorphic in a natural way to $G^* = T^n \times G_0$. Moreover, $G = G_1/H$ is isomorphic in a natural way to the factor group of $G^* = G_1/H_1$ by its finite subgroup $H^* = H/H_1$.

Let p^* be the natural homomorphism of G^* onto $G^*/H^* = G$. Using the fact that G is compact, and arguing in exactly the same way as in the first paragraph of the proof of Lemma 2, we find that G^* contains a compact subset K^* such that $p^*(K^*) = G$. But then it is easily seen that $G^* \subseteq K^*H^*$, so that since H^* is finite, G^* must be compact. Plainly, then, since $G^* = T^n \times G_0$, T^n must also be compact.

To complete the present proof, it only remains for us to show that T^n is isomorphic to the n-fold Cartesian product of circle groups: $T^n = C \times C \times \cdots \times C$. This we do as follows. Since $T^n = E^n/V_1$ is compact, it is impossible that V_1 should lie entirely in an $n - 1$ dimensional subspace of E^n. Hence V_1 contains n linearly independent vectors $v_1, \ldots, v_n$. Any vector $v \in E^n$ can be written in the form

$$v = t_1 v_1 + \cdots + t_n v_n.$$

We shall first prove that all the coefficients t_i must be rational. Suppose, to proceed by contradiction, that t_1 is irrational. Then, by forming $2v, 3v, \ldots$, and subtracting off suitable *integral* linear combinations of $v_1, \ldots, v_n$, we would find an infinite set of distinct elements of V_1, all having the form (1), where $0 \leqslant t_i < 1$, which, since V_1 contains only isolated points, is impossible. Thus we may write $t_i = p_i/q_i$, where p_i and q_i are integral; and using the same argument once more, the reader will verify without difficulty that, even if we allow v to range over all V_1, the denominators q_i are bounded below, so that each $v \in V_1$ may be written in the form (1) with $t_i = p_i/Q$, Q fixed once and for all, and p_i integral.

Let $u_1, \ldots, u_n$ be a second linearly independent set of vectors in V_1; write

$$u_i = \frac{1}{Q} \sum_{j=1}^{n} t_i^j v_j,$$

with t_i^j integral, and put

$$\det(u_1, \ldots, u_n; v_1, \ldots, v_n) = Q^{-n} |\det(t_i^j)|.$$

There evidently exists $u_1, \ldots, u_n$ for which $\det(u_1, \ldots, u_n; v_1, \ldots, v_n)$ takes on its smallest value, and then, as is easily verified, $\det(w_1, \ldots, w_n; u_1, \ldots, u_n)$ is a non-negative integer for each linearly independent set $w_1, \ldots, w_n$ of vectors in V_1. Let $w \in V_1$, and write $w = t_1 u_1 + \cdots + t_n u_n$. If $t_1 \neq 0$ and t_1 is not an integer, there exists $w' = t'_1 u_1 + \cdots + t'_n u_n$ with $0 \leqslant t'_i < 1$ and $t'_i \neq 0$. But then, putting $w_1 = w'$, $w_2 = u_2$, $w_3 = u_3, \ldots, w_n = u_n$, we have $\det(w_1, \ldots, w_n; u_1, \ldots, u_n) = t'_1 < 1$, a contradiction. It follows that every element $w \in V_1$ can be written as an integral linear combination of the vectors $u_1, \ldots, u_n$, so that V_1 is identical with the set of all integral linear combinations of these vectors. If we now use $u_1, \ldots, u_n$ as basis-vectors for E^n, we see easily $T^n = E^n/V_1$ is isomorphic to the n-fold Cartesian product of the group C of real numbers modulo integers with itself; and our proof is complete. Q.E.D.

Index

Ad λ 66
Adjoint of an operator 6
Adjoint representation 94
A(n) 58–59
Analytic functions of an operator 10–11
Automorphisms:
  Canonical factorization 181
  Diagonalization 198
  Jordan canonical form 194

Baker-Hausdorff theorem (*see* Campbell-Baker Hausdorff theorem)
Bilinear form 1

Campbell-Baker-Hausdorff theorem 68
Cartan criterion 98–99
Cartan subalgebra 95
  isomorphisms 138
Casimir operator 144
Chromosome 120
Classical algebras:
  models 123–125
  real forms 184–201
Classical groups 81–85
Closed subgroups 77
Commutative algebra 77
Compact groups 220–226
Complexification 87
Component of identity 38
Conjugations:
  definition 173
  equivalence 176
  positive 176
  in Algebras of type A_{n-1} 184–187
  in Algebras of type B_m, C_m, D_m 188–193
  in E_6 207–213
  in E_7 205–207
  in E_8 202–205
  in F_4 213–214
  in G_2 201–202
Connected group 37
Coordinates:
  definition 20
  for functions 21
  for tangent vectors 25
Covering group 42
Covering space 14–19

Derivative of a curve 29–30
Diagonalization 4
Differential equations 32
Dynkin diagram 120–121

Eigenspace 5
Eigenvalue:
  definition 5
  analytic dependence 10
Eigenvector 5
Exceptional algebras: 122
  models 126–136
  real forms 201–220
Exponential function 60

Factor algebra (*see* Quotient algebra)
Factor group (*see* Quotient group)
Fundamental group:
  definition 14
  of a covering space 16
  of a Lie group 42

Germ 44
GL(n) 58–59

Hermitian form 1
Hermitian operator 7
Homeomorphism 22
Homomorphism:
  of Lie algebras 55
    differentiability 64
    induced by group homomorphism 56
  of Lie groups 38
    defined by local homomorphism 39
    induced by algebra homomorphism 70
  of fundamental groups 14
Homotopy 13

Ideal:
  definitions 73, 88
  induced by invariant subgroups 74
Implicit function theorem 12, 28–29
Inner automorphism 138
Integrability conditions 56
Invariant subgroup 73
Invariant subspace 142
Invariant measure 19
Irreducible invariant subspace 143
Isomorphism:
  of Lie groups 38
    induced by algebra isomorphism 71
    induced by local isomorphism 41
    induced by mapping of roots 108
  of Lie algebras 55

Jacobian 27

Lie algebra 55, 86
Lie bracket·
  definition 50
  induced in a subalgebra 72
Lie group 37
Local homomorphism 38
Local isomorphism:
  definition 38
  inducing global isomorphism 41
  induced by algebra isomorphism 71
Local subgroup 44
Logarithmic coordinate system 64

Manifold 20

Natural bilinear form 94
Neighbors 114
Nilpotent algebra 88, 92
Non-singular bilinear form 2
Non-singular subspace 3

O(n) (Orthogonal group) 83

Positive definite form 4
Positive definite operator 8
Product manifold 32

Quotient algebra:
  definition 74
  determined by quotient group 75
Quotient group:
  definition 43, 75
  determining quotient algebra 75

Radical 160
Radical splitting theorem 160–163
Regular map 34
Representations:
  definition 87
  equivalence 146
Root space 94
Roots:
  definition 94
  positive 106
  simple set 105

Semi-direct product of groups 165
Semi-direct sum of algebras 159, 163
Semi-simple algebra:
  definition 93
  characterized by natural bilinear form 94
  characterized as a direct sum of simple algebras 99
  characterized by root structure 122
Signature 5
Simple algebra:
  definition 93
  characterized by root structure 115
  models 123–136
  real forms 215
Simply connected space 13, 16
Singular bilinear form 2
Skew-symmetric bilinear form 2

SL(n) (Special linear group) 81
Solvable algebra 88
Sp(n) (Symplectic group) 84
Special linear group (*see* SL(n))
Spectrum 5
Subalgebra 72
Subgroup:
  definition 44
  induced by subalgebra 72
Submanifold 34
Symmetric bilinear form 1
Symplectic group (*see* Sp(n))
Tangent vector:
  definition 23
  coordinates 25
  on submanifold 35
Topological group 19
Total commutativity 196

Unitary space 6

Vector field 31

Weight space 87, 90
Weight vector 87
Weights 87, 88, 146
Weyl group 140
Weyl reflection 140

WITHDRAWN

512.897
H376l
Hausner, Melvin
Lie groups, Lie algebras
47144

512.897
H376l
47144

Hausner, Melvin

Lie groups, Lie algebras

DAVIS MEMORIAL LIBRARY
3 7110 0003 1522 0

Notes on Mathematics and its Applications

General Editors: Jacob T. Schwartz, Courant Institute of Mathematical Sciences, and Maurice Lévy, Université de Paris

- *Algebraic Numbers and Algebraic Functions, E. Artin*
- *Collected Works of Hidehiko Yamabe, R. P. Boas*
- *A First Course in Functional Analysis, M. Davis*
- *Lectures on Modern Mathematics, M. Davis*
- *Singularities of Smooth Maps, J. Eells, Jnr.*
- *Advanced Ordinary Differential Equations, K. O. Friedrichs*
- *Special Topics in Fluid Dynamics, K. O. Friedrichs*
- *Lie Groups; Lie Algebras, M. Hausner and J. T. Schwartz*
- *Homotopy Theory and Duality, P. Hilton*
- *Lectures on Advanced Numerical Analysis, F. John*
- *Stability Techniques for Continuous Linear Systems, A. M. Krall*
- *An Introduction to Computer Programming, H. Mullish*
- *W* Algebras, J. T. Schwartz*
- *Exercises in Fortran, A. Silverman*
- *Nonlinear Elasticity, J. J. Stoker*

Additional volumes in preparation

GORDON AND BREACH
SCIENCE PUBLISHERS
150 FIFTH AVENUE
NEW YORK, N.Y. 1001